The Generalization of the R-transform for Invariant Pattern Representation

Thai V. Hoang[*,1,2] and Salvatore Tabbone[2]

[1] *MICA Center, HUST – CNRS/UMI 2954 – Grenoble INP, Hanoi, Vietnam*
[2] *LORIA, CNRS/UMR 7503, Nancy University, 54506 Vandoeuvre-les-Nancy, France*

Abstract

The beneficial properties of the Radon transform make it an useful intermediate representation for the extraction of invariant features from pattern images for the purpose of indexing/matching. This paper revisits the problem of Radon image utilization with a generic view on a popular Radon transform-based transform and pattern descriptor, the R-transform and R-signature, bringing in a class of transforms and descriptors spatially describing patterns at all directions and at different levels, while maintaining the beneficial properties of the conventional R-transform and R-signature. The domain of this class, which is delimited due to the existence of singularities and the effect of sampling/quantization and additive noise, is examined. Moreover, the ability of the generic R-transform to encode the dominant directions of pattern is also discussed, adding to the robustness to additive noise of the generic R-signature. The stability of dominant direction encoding by the generic R-transform and the superiority of the generic R-signature over existing invariant pattern descriptors on grayscale and binary noisy datasets have been confirmed by experiments.

1 Introduction

Many descriptors have been proposed in literature for the extraction of patterns' invariant features [1, 2] using techniques that allow invariance to rotation, scaling, translation (RST) or their combinations. Translation and scaling invariance could be obtained by using the Fourier [3] and Mellin [4] transforms respectively; rotation invariance by computing the harmonic expansion [5] or performing the discrete Fourier transform on the circular coordinate of the pattern image in polar space [6], etc. However, the task of combining several techniques to have all RST while guaranteeing the discriminatory power of the invariant features is challenging and has attracted attention of many researchers. Most of the existing methods do not allow invariance to all RST, they usually require normalizations for the unavailability of any of RST. For example, methods based on the theory of moments [7] usually normalize input pattern images regarding their centroid position and size: the pattern's centroid is required to coincide with the origin of the coordinate system and the longest distance between this centroid and a pattern's point is equal to one. These normalizations usually introduce errors, are sensitive to noise, and thus induce inaccuracy in a later matching process.

Radon transform-based methods are different from the others in the sense that the Radon transform of pattern images is used as an intermediate representation upon which invariant features are extracted from for the purpose of indexing/matching. There are some reasons for the utilization of Radon

[*] Corresponding author. Tel: +33 3 54 95 85 60; Fax: +33 3 83 27 83 19.
Bureau B226, LORIA, Campus Scientifique, BP 239, 54506 Vandoeuvre-les-Nancy Cedex, France.
E-mail address: vanthai.hoang@loria.fr (Thai V. Hoang), tabbone@loria.fr (Salvatore Tabbone).

transform. Firstly, it is a rich transform with one-to-many mapping, each pattern's point lies on a set of lines in the spatial pattern image and contributes a curve to the Radon image. Secondly, it is a lossless transform, pattern images can even be reconstructed accurately by the inverse Radon transform. Thirdly, it has low complexity, requiring $O(N \log N)$ operations for an input image of N pixels [8]. And finally and more importantly, it has useful properties concerning RST transformations which have been applied on a pattern image. By applying the Radon transform on an RST transformed pattern image, the transformation parameters are encoded in the radial (for translation and scaling) and angular (for rotation) coordinates of the obtained Radon image [9]. Current techniques thus usually exploit this encoded information to define invariant features.

A pioneer work in this direction is the R-transform, which gives rise to the R-signature [10], obtained by using an integral function and then the discrete Fourier transform for the radial and angular coordinates of the Radon image respectively. Similarly, the Φ-signature [11] is computed by using an integral function along the angular coordinate of the Radon image to get rotation invariance. Invariance to translation and scaling is made possible by normalizations. The strength of these two approaches is simplicity, however, the obtained signatures have low discriminatory power as there is a loss of information in the compression process from the Radon image to 1D signatures. Moreover, the required normalizations for the Φ-signature prevent it from being applied to noisy images.

There was an effort to apply the 2D Fourier–Mellin transform on the Radon image [12] to get invariance to rotation and scaling. In this approach, Mellin transform and harmonic expansion are applied on the radial and angular coordinates of the Radon image respectively. The main weakness of this approach is the lack of translation invariance. This drawback has been overcome in [13] by replacing the Mellin transform by the 1D Fourier–Mellin transform for the radial coordinate. Recently, a set of spectral and structural features has also been extracted from the Radon image for pattern description [14]. In this set, the "degree of uniformity" is essentially the R-transform and the "longest line" is the information encoded in the generic R-signature described in this paper. However, and more importantly, this set of features is not invariant to rotation and consequently, in the matching step, these features need to be rotated to all possible rotating angles corresponding to potential pattern's orientations in order to compute patterns' similarity. Long matching time may prevent the application of this approach in real systems.

Another direction in using the Radon transform for pattern description is to extract pattern features directly from the Radon image, similar to the way the Hough transform [15] is used. For example, pattern primitives in edge form are detected from the Radon image and represented analytically in [16]. Moreover, their spatial relations can be made explicit [17] and this leads to a taxonomy of patterns for their characterization [18]. This approach, however, is quite limited as it requires that the edge primitives have analytical form. Generalizations of the Radon transform, called the trace and geometric transforms, have also been proposed and used for image description [19, 20] by using functionals other than integral and by extending the functional domain from a line to a region delimited by a closed contour. However, their application is restricted due to high computational complexity.

Among the Radon transform-based pattern descriptors, R-signature is the most popular due to its simplicity and has been successfully applied to several applications (e.g., symbol recognition [21], activity recognition [22, 23], and orientation estimation [24]). This paper provides a generic view on the R-transform and the R-signature while maintaining their beneficial properties, leading to three main theoretical contributions. The first is a better understanding of the discriminatory power of the generic R-signature, which results from the exploitation of variation in the accumulation of the pattern image along all the parallel lines, leading to an increase in performance. The second is a discussion on the reasonable range of the generalization, which is limited due to the existence of singularities

and the sensitivity to sampling/quantization and additive noise. The last is the ability of the generic R-transform to represent dominant directions of patterns, even in the present of noise, resulting in the superiority of the generic R-signature on noisy datasets over comparison methods. A preliminary study of this generalization has been published in [25].

The remainder of this paper is organized as follows. Section 2 gives some background on the Radon transform along with a brief review of the conventional R-transform and R-signature. Section 3 presents the generalization of the R-transform and R-signature along with their geometric interpretation and their properties. A discussion on the meaningful domain of these generic transform/signature is carried out in Section 4. Theoretical arguments on their robustness to additive noise and their ability to encode dominant directions of patterns are presented in Section 5. Experimental results are given in Section 6 and finally conclusions are drawn in Section 7.

2 Basic material

This section provides some basics of the Radon transform, its definition and its derived beneficial properties. The inspiration for the proposal of the R-signature from these properties will also be presented afterwards.

2.1 The Radon transform

Let $f(x, y) \in \mathbb{R}^2$ be a 2D function and $L(\theta, \rho)$ be a straight line in $\mathbb{R}^2$ represented by

$$L = \{(x, y) \in \mathbb{R}^2 : x \cos\theta + y \sin\theta = \rho\},$$

where θ is the angle L makes with the y axis and ρ is the distance from the origin to L. The Radon transform [26] of f, denoted by $\mathcal{R}_f$, is a functional defined on the space of lines $L(\theta, \rho)$ by the line integral along each line:

$$\mathcal{R}_f(L) = \mathcal{R}_f(\theta, \rho) = \int_L f(x, y)\, \mathrm{d}x\mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, \delta(\rho - x\cos\theta - y\sin\theta)\, \mathrm{d}x\mathrm{d}y. \tag{1}$$

In the field of shape analysis and recognition, the function $f(x, y)$ is constrained to the following particular definition:

$$f(x, y) = \begin{cases} 1 & \text{if } x \in D \\ 0 & \text{otherwise,} \end{cases}$$

where D is the domain of the binary shape represented by $f(x, y)$. In the illustration of the Radon transform in Fig. 1, the shaded region represents the region D. The value of the line integral in Eq. (1) is equal to the length of the intersection between the line L and the shaded region.

The Radon transform has some properties that are beneficial for invariant pattern recognition problems as outlined below:

$P\mathbf{1}$ *linearity*: The Radon transform is linear.

$$\mathcal{R}_{(f+g)}(\theta, \rho) = \mathcal{R}_f(\theta, \rho) + \mathcal{R}_g(\theta, \rho).$$

$P\mathbf{2}$ *periodicity*: The Radon transform of $f(x, y)$ is periodic in the variable θ with period 2π.

$$\mathcal{R}_f(\theta, \rho) = \mathcal{R}_f(\theta + 2k\pi, \rho), \qquad \forall k \in \mathbb{Z}.$$

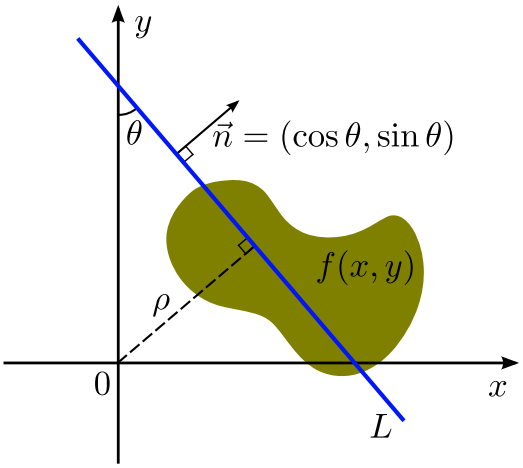


Figure 1: Graphical illustration of the Radon transform of a function $f(x, y)$. The Radon transform is a mapping from the image space (x, y) to the parameter space (θ, ρ) and can be mathematically represented by a line integral of $f(x, y)$ along all the lines L parameterized by (θ, ρ) represented in the image space (x, y).

*P***3** *semi-symmetry*: The Radon transform of $f(x, y)$ is semi-symmetric.

$$\mathcal{R}_f(\theta, \rho) = \mathcal{R}_f(\theta \pm \pi, -\rho).$$

*P***4** *translation*: A translation of $f(x, y)$ by a vector $\vec{u} = (x_0, y_0)$ results in a shift in the variable ρ of $\mathcal{R}_f(\theta, \rho)$ by a distance $d = x_0 \cos\theta + y_0 \sin\theta$ that is equal to the length of the projection of $\vec{u}$ onto the line $x \cos\theta + y \sin\theta = \rho$.

$$\mathcal{R}_f(\theta, \rho) \rightarrow \mathcal{R}_f(\theta, \rho - x_0 \cos\theta - y_0 \sin\theta).$$

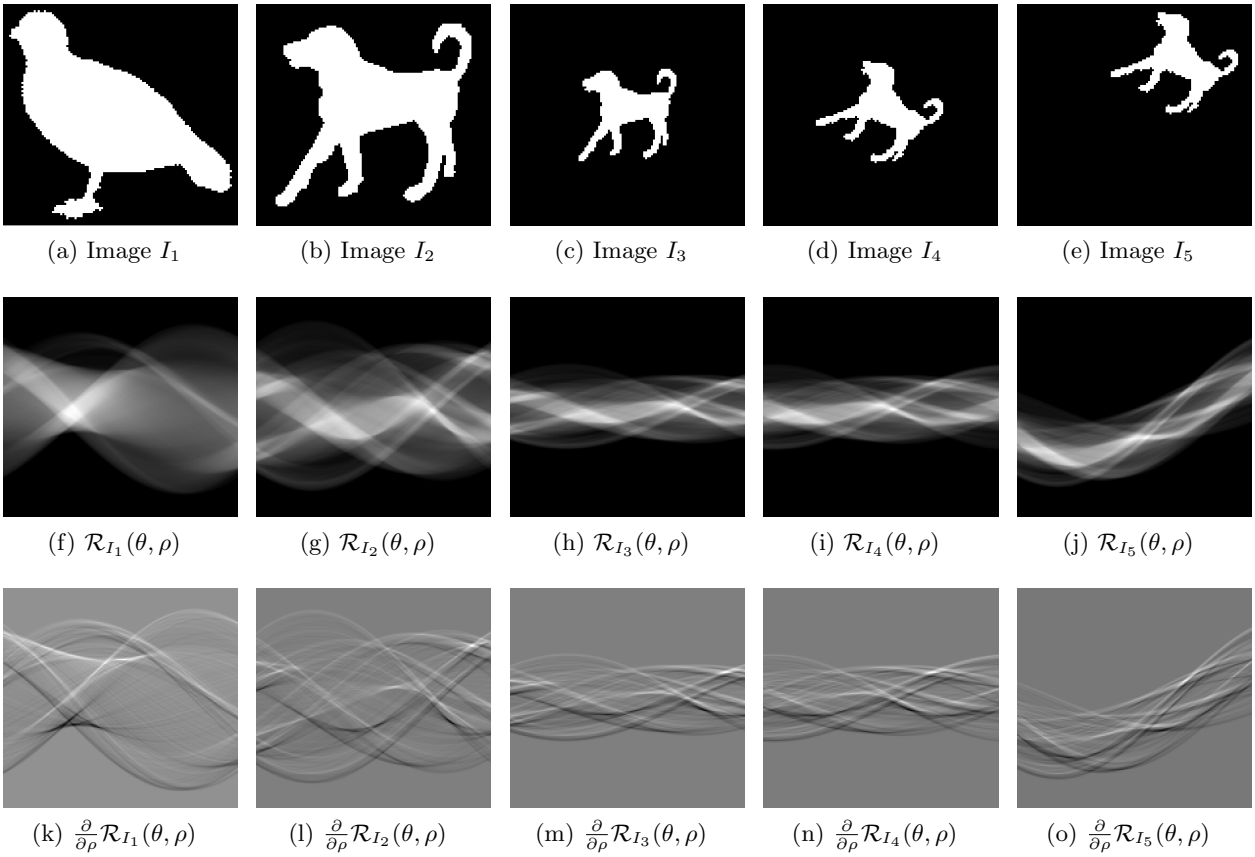
*P***5** *rotation*: A rotation of $f(x, y)$ by an angle θ_0 implies a circular shift in the variable θ of $\mathcal{R}_f(\theta, \rho)$ by a distance θ_0.

$$\mathcal{R}_f(\theta, \rho) \rightarrow \mathcal{R}_f(\theta + \theta_0, \rho).$$

*P***6** *scaling*: A scaling of $f(x, y)$ by a factor α results in a scaling in the variable ρ and the amplitude of $\mathcal{R}_f(\theta, \rho)$ by the factors α and $\frac{1}{\alpha}$ respectively.

$$\mathcal{R}_f(\theta, \rho) \rightarrow \frac{1}{\alpha}\mathcal{R}_f(\theta, \alpha\rho).$$

Thus, by applying the Radon transform on an RST transformed pattern image, the RST transformation parameters are encoded in the radial (for translation and scaling) and angular (for rotation) coordinates of the obtained Radon image [9]. Current techniques usually exploit this encoded information to define invariant descriptors. Fig. 2 illustrates the effect of RST transformations on the Radon image. The top row contains two original pattern images I_1, I_2 (Figs. 2a, 2b) and the RST transformed versions I_3, I_4, I_5 (Figs. 2c–2e) of I_2. The second row shows the Radon transform of these pattern images. It is observed that the Radon transforms of I_1 and I_2 are totally different while there exists resemblance among that of I_2–I_5 as the result of properties $P4$–$P6$ of the Radon transform: scaling ($I_2 \rightarrow I_3$) becomes a compression in constant-θ slices, rotation ($I_3 \rightarrow I_4$) becomes a constant shift in the angular direction, and translation ($I_4 \rightarrow I_5$) becomes a sinusoidal shift in the radial direction.



(a) Image I_1 (b) Image I_2 (c) Image I_3 (d) Image I_4 (e) Image I_5

(f) $\mathcal{R}_{I_1}(\theta,\rho)$ (g) $\mathcal{R}_{I_2}(\theta,\rho)$ (h) $\mathcal{R}_{I_3}(\theta,\rho)$ (i) $\mathcal{R}_{I_4}(\theta,\rho)$ (j) $\mathcal{R}_{I_5}(\theta,\rho)$

(k) $\frac{\partial}{\partial\rho}\mathcal{R}_{I_1}(\theta,\rho)$ (l) $\frac{\partial}{\partial\rho}\mathcal{R}_{I_2}(\theta,\rho)$ (m) $\frac{\partial}{\partial\rho}\mathcal{R}_{I_3}(\theta,\rho)$ (n) $\frac{\partial}{\partial\rho}\mathcal{R}_{I_4}(\theta,\rho)$ (o) $\frac{\partial}{\partial\rho}\mathcal{R}_{I_5}(\theta,\rho)$

Figure 2: Illustration of the properties of the Radon transform performed on pattern images and the differentiation operator. The first row contains two original pattern images (a)–(b) and the RST transformed versions (c)–(e) of the pattern image in (b). The second row shows the Radon transform of these pattern images and the third row shows the partial derivative of these Radon images with respect to the variable ρ. The intensity of these images has been rescaled to fit the display range.

2.2 The R-signature

The conventional R-transform of a 2D function $f(x,y) \in \mathbb{R}^2$ proposed in [10] has the following definition:

$$R_{f2}(\theta) = \int_{-\infty}^{\infty} \mathcal{R}_f^2(\theta,\rho)\,\mathrm{d}\rho. \tag{2}$$

The integration computed along constant-θ slices of the Radon image in Eq. (2) makes $R_{f2}(\theta)$ invariant to translation and scaling, except for a multiplicative factor $\frac{1}{\alpha^3}$ resulting from the scaling factor α in $f(x,y)$, and periodic with period π. Furthermore, in order to have a representation totally invariant to RST transformations, the authors of [10] have proposed to use the magnitude of the discrete Fourier transform performed on the discretized $R_{f2}(\theta_n)$ normalized by the DC component as

$$FR_{f2}(k) = \left| \frac{\sum_{n=0}^{N-1} R_{f2}(\theta_n)\,\mathrm{e}^{-\frac{2\pi i}{N}kn}}{\sum_{n=0}^{N-1} R_{f2}(\theta_n)} \right|, \qquad k = 0,\ldots,N-1, \tag{3}$$

and the conventional R-signature of $f(x,y)$ is originally defined as

$$\left[FR_{f2}(1), FR_{f2}(2), \ldots, FR_{f2}(N-1)\right].$$

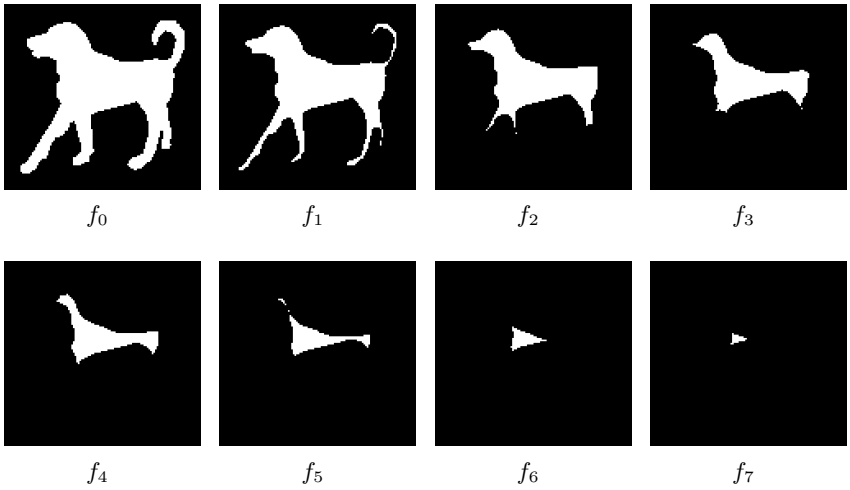


Figure 3: Eight images obtained by segmenting the distance transform of the shape image I_2 in Fig. 2b at 8 equi-distant levels. The R-transform has been computed on these images, $R_{f_i 2}(\theta)$ with $i = 0 \rightarrow 7$, in order to increase the discriminatory power for shape recognition/matching applications.

Other solutions for a totally invariant representation based on the R-transform could be obtained by centering $R_{f2}(\xi)$ using the forward and then reverse discrete Fourier transforms as

$$IFR_{f2}(\theta_n) = \frac{\sum_{k=0}^{N-1} FR_{f2}(k)\, \mathrm{e}^{\frac{2\pi i}{N} kn}}{\sum_{k=0}^{N-1} FR_{f2}(k)}, \qquad n = 0, \ldots, N-1. \tag{4}$$

or by canceling the phase by multiplying appropriate harmonic expansions as in [27]. It is not difficult to prove that $FR_{f2}(\xi)$ and $IFR_{f2}(\theta)$ are invariant to RST transformations. For the rest of this paper, $FR_{f2}(\xi)$ will be used as the basic definition of the R-signature.

3 Generalization of the R-signature

The R-signature defined in Subsection 2.2, originally proposed for invariant shape representation, has been extended in [10] by computing $FR_{f_i 2}(\theta)$, where f_i $(i = 0 \rightarrow 7)$ are derived from a shape image f by segmenting its distance transform [28] at 8 equi-distant levels. This extension leads to an increase in the discriminatory power of the R-signature because the derived shape images f_i preserve the topology of f and, when i increases, the level of deformation decreases. The derived images obtained by segmenting the distance transform of the shape image I_2 in Fig. 2b is illustrated in Fig. 3. This extension, however, works only with silhouette shapes and it is difficult to use it with noisy pattern images.

Another extension, which is orthogonal to the extension described above, is proposed in this section by generalizing the R-transform in Eq. (2) to further increase the discriminatory power. It will be shown that the R-transform in Eq. (2) is just a special case of a class of transforms sharing beneficial properties for pattern representation and matching. This section provides the definition of this class of transforms and their geometric interpretation in the spatial domain. The properties of the signatures defined based on these transforms are given afterwards.

3.1 Definition

The basic idea of the R-transform is the use of an integration to overcome the scaling and translation problems that remain in constant-θ slices of the Radon image (properties $P4$ and $P6$). For any function

$g(x)$ and its transformed version $h(x) = g(\alpha x - x_0)$, their integration:

$$\int_{-\infty}^{\infty} h(x)\,\mathrm{d}x = \frac{1}{\alpha}\int_{-\infty}^{\infty} g(x)\,\mathrm{d}x$$

differs by a multiplicative factor depending solely on the scaling parameter α, which could be easily removed by normalization in a later processing step. This strategy could actually be extended to any operator $\mathcal{O}_1$ satisfying:

$$\mathcal{O}_1(h(\cdot)) = \kappa_1(x_0, \alpha)\,\mathcal{O}_1(g(\cdot)), \tag{5}$$

where $\kappa_1(x_0, \alpha)$ is a function which depends solely on x_0 and α. Possible other choices for $\mathcal{O}_1$ could be the 1D Fourier–Mellin transform [13] and some measures like maxima, median, etc.

In addition to the integral operator, a square operator is also employed in the definition of the R-transform to avoid the singularities (discussed in Section 4) while preserving properties $P4$ and $P6$ of the Radon transform. This operator in turn could also be replaced by any operator $\mathcal{O}_2$ satisfying:

$$\mathcal{O}_2(h(x)) = \kappa_2(x_0, \alpha)\,\mathcal{O}_2\left(\frac{1}{\alpha}g(\alpha x - x_0)\right), \tag{6}$$

where $\kappa_2(x_0, \alpha)$ is a function which depends solely on x_0 and α. Some operators like exponentiation, differentiation, etc. could be used for $\mathcal{O}_2$. As an example, the bottom row of Fig. 2 shows the results obtained by using differentiation with respect to the variable ρ of the Radon image. It is clear that differentiation retains properties $P4$ and $P6$ of the Radon transform, however, it accentuates small variation in the Radon image due to sampling/quantization and additive noise.

If there exists two operators $\mathcal{O}_1$ and $\mathcal{O}_2$ satisfying Eqs. (5) and (6) respectively, the combined operator $\mathcal{O}_{12} = \mathcal{O}_1 \circ \mathcal{O}_2$, when applied on constant-θ slices of the Radon image, will overcome the scaling and translation problems. Despite the existence of several choices for $\mathcal{O}_1$ and $\mathcal{O}_2$, the generalization of the R-transform described in this section uses an integration for $\mathcal{O}_1$ and an exponentiation for $\mathcal{O}_2$. These choices of operators result in a generic transform that has many beneficial properties and superior performance over existing methods, which will be described and demonstrated in the following sections. Thus, for a 2D function $f(x, y) \in \mathbb{R}^2$ and $m \in \mathbb{R}$, the generic R-transform of $f(x, y)$, denoted as $R_{fm}(\theta)$, is defined as

$$R_{fm}(\theta) = \int_{-\infty}^{\infty} \mathcal{R}_f^m(\theta, \rho)\,\mathrm{d}\rho. \tag{7}$$

Evidently, by setting $m = 2$, $R_{fm}(\theta)$ in Eq. (7) becomes $R_{f2}(\theta)$ in Eq. (2). The utilization of the exponent m as a parameter makes $R_{fm}(\theta)$ a generic version of $R_{f2}(\theta)$. Furthermore, by varying the value of m, a class of transforms could be obtained and this in turn results in a class of signatures. The derivation of the generic R-signature, $FR_{fm}(\xi)$, from $R_{fm}(\theta)$ follows strictly the process described in Subsection 2.2.

3.2 Geometric interpretation

Recall that the value of $\mathcal{R}_f(\theta, \rho)$ is the result of a line integral of $f(x, y)$ along the line $L(\theta, \rho)$ parameterized by (θ, ρ). Consequently, the generic R-transform defined based on $\mathcal{R}_f(\theta, \rho)$ by computing integrals over the variable ρ has some geometric interpretations as follows.

The generic R-transform of f, $R_{fm}(\theta)$, in Eq. (7) is basically an integral of $\mathcal{R}_f^m(\theta, \rho)$ computed over the variable ρ of the Radon image. In other words, this integral is computed by using the result of line integral along all the lines parameterized by a fixed value of θ and different values of ρ. Sharing the same

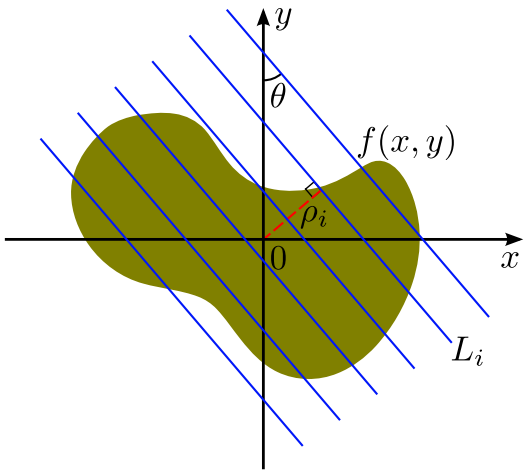


Figure 4: Geometric illustration of the generic R-transform of a function $f(x, y)$. Shown in the figure is the set of all the lines $L(\theta, \cdot)$ used to compute the value of $R_{fm}(\theta)$ the formula in Eq. (7). $R_{fm}(\theta)$ contains the encoded spatial information of the pattern image in the direction making an angle θ with the y axis.

value of θ means that these lines are parallel in the spatial space (as depicted in Fig. 4) and $R_{fm}(\theta)$ encodes the spatial information of the pattern image $f(x, y)$ in the direction making an angle θ with the y axis. Encoding $f(x, y)$ at different directions is possible by varying θ to have $R_{fm}(\cdot)$ and $R_{fm}(\cdot)$ could be then interpreted as containing the encoded spatial information of $f(x, y)$ at all directions.

The role of the exponent m in Eq. (7), besides setting up a class of transforms, is to make $R_{fm}(\cdot)$ discriminatory at different values of m by exploiting the variation in $\mathcal{R}_f(\theta, \cdot)$, which in turn is the variation in the length of the intersection of $f(x, y)$ with all paralleling lines L_i making an angle θ with the y axis. Evidently, at $m = 2$, $R_{fm}(\cdot)$ has the same interpretation and discriminatory power as that of the conventional R-transform. The interestingness lies in large m at which $R_{fm}(\cdot)$ has the capability to encode the dominant direction or "longest line" as will be demonstrated in Section 5. In addition, due to the singularity at $m = 1$ (Subsection 4.1), it is anticipated that the generic R-transform will have a higher discriminatory power when the value of m goes far way from 1. Inversely, it should be noted also that, when $m < 1$, $R_{fm}(\cdot)$ weights more on shorter lines.

3.3 Properties

The generic R-transform as defined in Eq. (7) has some beneficial properties as follows (a proof sketch is given in Appendix A):

- *Periodicity*: The generic R-transform of $f(x, y)$ is periodic in the θ coordinate with period π.

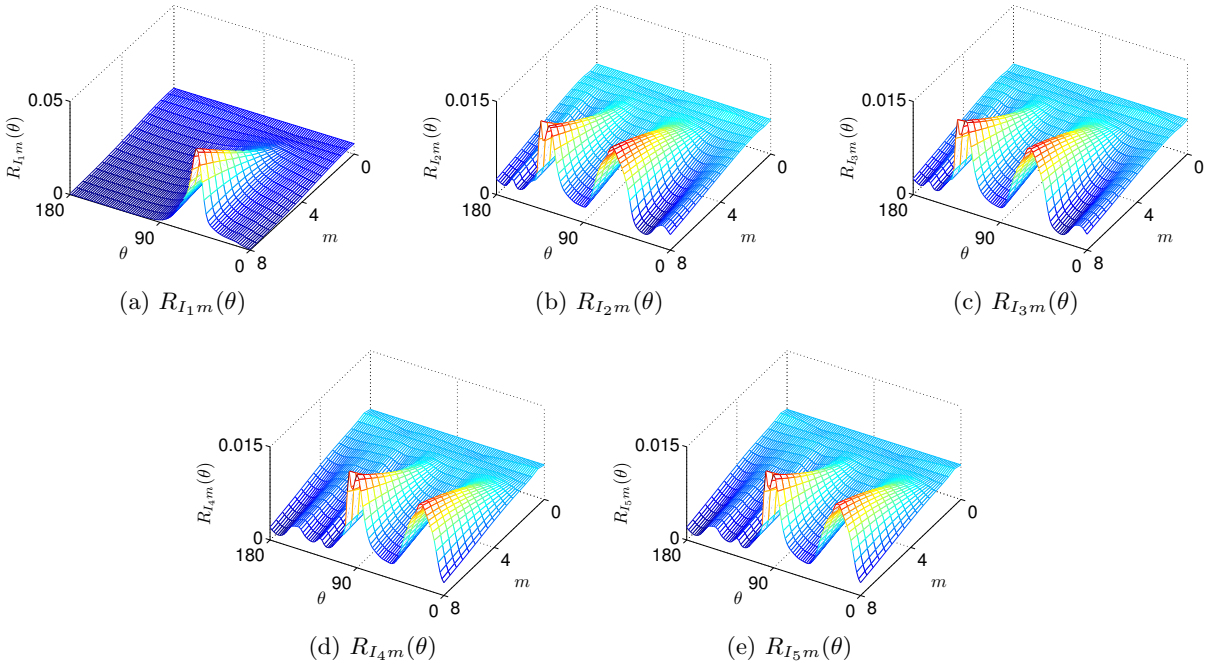
$$R_{fm}(\theta) = R_{fm}(\theta + k\pi), \qquad \forall k \in \mathbb{Z}.$$

- *Translation*: The generic R-transform of $f(x, y)$ is invariant to translation.

$$R_{f'm}(\theta) = R_{fm}(\theta).$$

- *Rotation*: A rotation of $f(x, y)$ by an angle θ_0 implies a circular shift of $R_{fm}(\theta)$ by a distance θ_0.

$$R_{f'm}(\theta) = R_{fm}(\theta + \theta_0).$$



(a) $R_{I_1m}(\theta)$ (b) $R_{I_2m}(\theta)$ (c) $R_{I_3m}(\theta)$

(d) $R_{I_4m}(\theta)$ (e) $R_{I_5m}(\theta)$

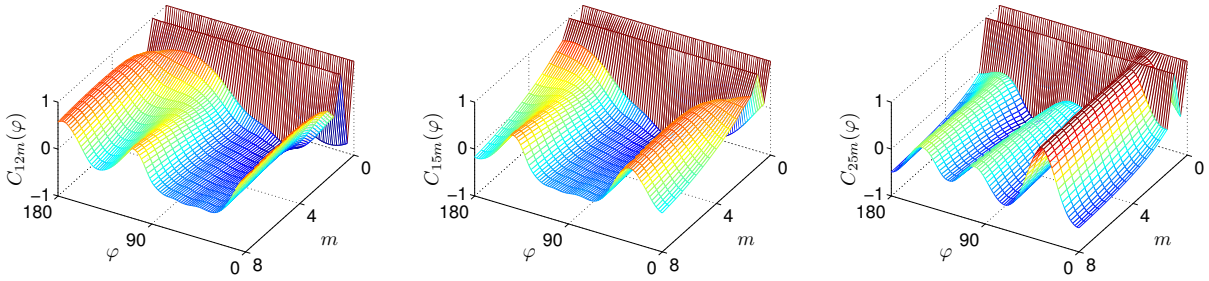
Figure 5: Illustration of the properties of the generic R-transform by using pattern images in the top row of Fig. 2 with the exponent $m = 0 \rightarrow 6$. The generic R-transform is invariant to translation and scaling and converts rotation in the spatial pattern image to translation in $R_{I_km}(\cdot)$ by a distance equal to the rotating angle.

- *Scaling*: A scaling of $f(x, y)$ by a factor α results in a scaling in the amplitude of $R_{fm}(\theta)$ by a factor $\frac{1}{\alpha^{m+1}}$.

$$R_{f'm}(\theta) = \frac{1}{\alpha^{m+1}} R_{fm}(\theta). \tag{8}$$

From these properties, it is straightforward that the generic R-signature of $f(x, y)$, $FR_{fm}(\theta)$ or $IFR_{fm}(\theta)$, defined based on the generic R-transform of $f(x, y)$, $R_{fm}(\theta)$, as in Eqs. (3) or (4) in Subsection 2.2 is totally invariant to RST transformations. Illustration of the properties concerning RST transformations of the generic R-transform is given in Fig. 5 using pattern images in the top row of Fig. 2. The value of $R_{I_km}(\cdot)$ with $k = 1 \rightarrow 5$ has been normalized by the area they make with the θ axis, $\int_0^\pi R_{I_km}(\theta)\, d\theta$, for better viewing. The two images I_1, I_2 in Fig. 2a, 2b are not similar and as a consequence $R_{I_1m}(\cdot)$, $R_{I_2m}(\cdot)$ in Fig. 5a, 5b have different pattern. The images I_3, I_4, and I_5 in Fig. 2c–2e are transformed versions of the image I_2 in Fig. 2b then $R_{I_3m}(\cdot)$, $R_{I_4m}(\cdot)$, and $R_{I_5m}(\cdot)$ in Fig. 5c–5e have the same pattern with $R_{I_2m}(\cdot)$ in Fig. 5b. Furthermore, it is evident that the generic R-transform is invariant to scaling ($R_{I_2m}(\cdot) \simeq R_{I_3m}(\cdot)$) and translation ($R_{I_4m}(\cdot) \simeq R_{I_5m}(\cdot)$). It converts rotation in the spatial pattern image I_k into a circular shift in the variable θ of $R_{I_km}(\cdot)$ by a distance equal to the rotating angle ($R_{I_3m}(\cdot) \rightarrow R_{I_4m}(\cdot)$).

A quantitative evaluation of the invariant properties of the generic R-transform is given in Fig. 6 using the normalized cross correlation between the three possible pairs of the generic R-transforms $R_{I_1m}(\theta)$, $R_{I_2m}(\theta)$, and $R_{I_5m}(\theta)$ from Fig. 5. Normalized cross correlation is selected for the sake of overcoming the constant multiplicative factor $\frac{1}{\alpha^{m+1}}$ in Eq. (8) and the remaining rotation. To overcome the remaining rotating parameter, at a specific value of m, the correlation is calculated for all possible rotating angles, meaning that one of the two generic R-transforms is circular shifted along its θ direction



(a) $C_{12m}(\varphi)$ (b) $C_{15m}(\varphi)$ (c) $C_{25m}(\varphi)$

Figure 6: The normalized cross correlation between the three possible pairs of the generic R-transforms $R_{I_1m}(\theta)$, $R_{I_2m}(\theta)$, and $R_{I_5m}(\theta)$ from Fig. 5. For each pair of generic R-transforms and at a specific value of m, 180 correlation values are calculated after circular shifting one of the two generic R-transforms along its θ direction by 180 possible values from 0 to 179 with increment of 1.

by 180 possible values from 0 to 179 with increment of 1 before computing the correlation. Denoting φ as the shifting distance, the correlation between $R_{I_im}(\theta)$ and $R_{I_jm}(\theta)$ at φ is defined as

$$C_{ijm}(\varphi) = \text{corr}\left(R_{I_im}(\theta), R_{I_jm}(\theta+\varphi)\right),$$

where $\text{corr}(A, B)$ is the normalized cross correlation function between two input vectors A and B of length n calculated using the following formula:

$$\text{corr}(A, B) = \frac{\sum_{i=1}^{n}(A_i - \bar{A})(B_i - \bar{B})}{\sqrt{\left(\sum_{i=1}^{n}(A_i - \bar{A})^2\right)\left(\sum_{i=1}^{n}(B_i - \bar{B})^2\right)}},$$

where $\bar{A}$ and $\bar{B}$ are the mean values of A and B respectively.

The three surfaces $C_{ij}(\varphi)$ corresponding to the three possible pairs of three generic R-transforms have some distinct characteristics. Firstly, they all have two constant bars at $m = 0, 1$ due to the singularities of the generic R-transform at $m = 0, 1$, which will be discussed in Subsection 4.1. Secondly, at a specific value of m, the maximum value of $C_{25m}(\varphi)$ is almost 1 while that of $C_{12m}(\varphi)$ and $C_{15m}(\varphi)$ is always less than 0.67, or more generally $C_{ijm}(\varphi)$ is peaky only when the two pattern images I_i and I_j are similar. The non-peaky and peaky maxima exhibit the discriminatory power of the proposed descriptor and the maximum of nearly 1 means, in this case, that the generic R-transform is invariant to translation and scaling. Moreover, the value of φ^* corresponding to the peak in $C_{ijm}(\varphi)$ denotes the difference in orientation (in degree) between the patterns in I_i and I_j.

3.4 Implementation

From the definition of the generic R-signature in Subsection 3.1, its calculation could be separated into three steps: Radon, generic R-transform, and discrete Fourier transform. The Radon transform is computed based on recursively defined digital straight lines [29, 30], requiring $O(N \log N)$ operations for an images of $N = n \times n$ pixels. Similarly, for 1D digital data of M samples, the discrete Fourier transform can be implemented using the FFT algorithm requiring $O(M \log M)$ operations and the remaining generic R-transform requires $O(M)$ operations.

Apparently, there is no increase in the computational complexity when generalizing the R-signature. The generic R-signature maintains the simplicity of the conventional R-signature proposed in [10], leading to a simple and reasonably fast computation. Compared to the Radon transform-based R2DFM [12] and RFM [13] descriptors, the computation of the generic R-signature requires the least amount of time.

4 The domain for the exponent m

The generic R-transform, as defined in Eq. (7), theoretically produces a class of transforms having an infinite number of members obtained by varying the value of the exponent m. However, in reality, the domain for reasonable values of m is limited, not all the space $\mathbb{R}$, due to the existence of singularities and the sensitivity to sampling/quantization and additive noise.

4.1 Singularities

The generic R-transform has two singularities at $m = 0$ and $m = 1$:

$$R_{f0}(\theta) = \int_{-\infty}^{\infty} \mathcal{R}_f^0(\theta, \rho)\, \mathrm{d}\rho = \rho_{\max} - \rho_{\min} = \text{const},$$

$$R_{f1}(\theta) = \int_{-\infty}^{\infty} \mathcal{R}_f^1(\theta, \rho)\, \mathrm{d}\rho = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, \mathrm{d}x\mathrm{d}y = m_{00} = \text{const},$$

where m_{00} is the zeroth-order moment of $f(x, y)$. $R_{f0}(\cdot)$ and $R_{f1}(\cdot)$ hence contain no discriminatory information about $f(x, y)$, except for scaling when $m = 1$, and they should not be used to represent patterns for the purpose of recognition and/or matching. Additionally, when m reaches $+\infty$, as the cumulative sum of $f(x, y)$ along a line $L(\theta, \rho)$ is most of the time greater than 1, then

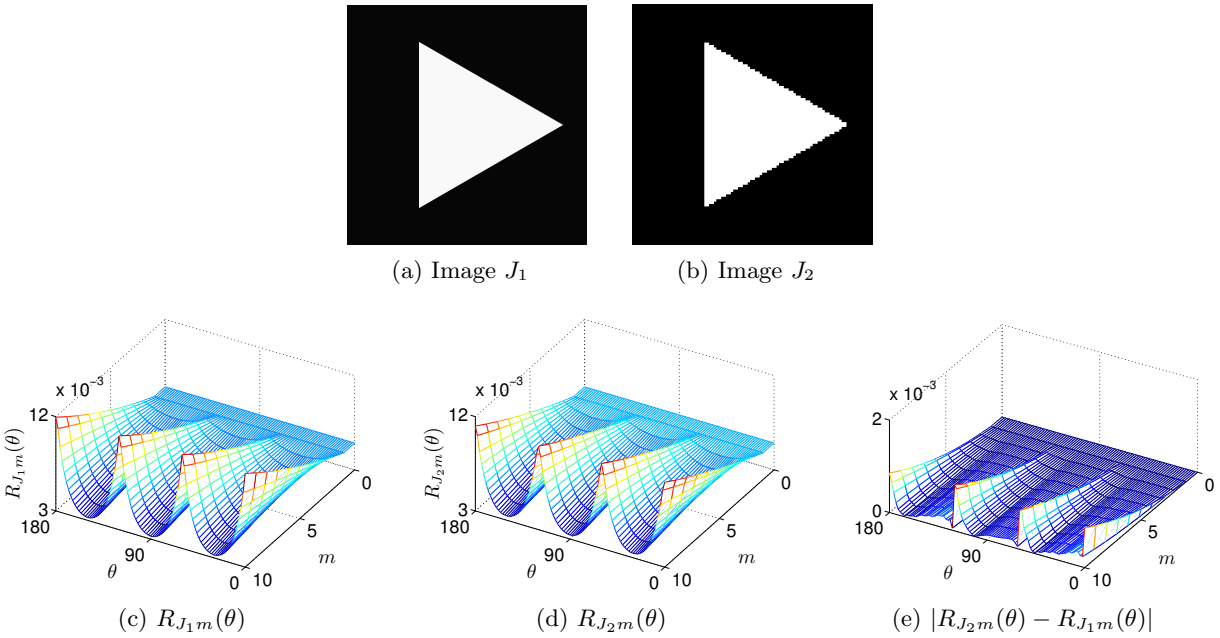
$$\lim_{m\to+\infty} R_{fm}(\theta) = \lim_{m\to+\infty} \int_{-\infty}^{\infty} \mathcal{R}_f^m(\theta, \rho)\, \mathrm{d}\rho = +\infty. \tag{9}$$

Even though $m = \infty$ has no practical meaning, the result in Eq. (9) implies that in practice m cannot have excessive large values. Furthermore, negative value of m should also be avoided due to the sensitivity of negative power function to very small values. More precisely, at the furthest point from the pattern centroid's position, the intersection between the tangent line $L(\theta^\star, \rho^\star)$ and the pattern image $f(x, y)$ has infinitesimal length, inducing a very small value at $\mathcal{R}_f(\theta^\star, \rho^\star)$. Taking negative power of this value produces a very large value and is sometimes out of the representing capability of digital computers.

4.2 Sensitivity to sampling/quantization and additive noise

By definition, the Radon transform is essentially the projection of the spatial pattern image $f(x, y)$ along all the lines $L(\theta, \rho)$ and, due to this projection, the Radon transform has the ability to suppress variation in the pattern images [31]. However, as the generic R-transform is defined based on the exponentiation of $\mathcal{R}_f(\theta, \rho)$, the remaining variation due to noise in $\mathcal{R}_f(\theta, \rho)$ will result in variation in $R_{fm}(\cdot)$ at different levels according to the value of the exponent m. A too high value of m will cause high variation in $R_{fm}(\cdot)$ and make it very different from the ideal analytical values. The heavily deformed $R_{fm}(\cdot)$ due to noise will make the representation inappropriate for recognition/matching.

Sampling and quantization could be considered as processes that add noise to the original analytical pattern images. In this sense, the pattern images processed by digital computers are noisy images and the variation in $\mathcal{R}_f(\theta, \rho)$ is unavoidable. Fig. 7 illustrates the sensitivity of the generic R-transform to sampling and quantization. The image J_2 in Fig. 7b is a sampled and quantized version of an analytical triangle J_1 in Fig. 7a and has a size of 100×100 pixels. The value of $R_{J_k m}(\cdot)$ with $k = 1, 2$ in Figs. 7c–7d has been normalized by the area they make with the θ axis, $\int_0^\pi R_{J_k m}(\theta)\, d\theta$, for better viewing. The difference in the normalized values of the generic R-transform of J_1 and J_2, $|R_{J_2 m}(\theta) - R_{J_1 m}(\theta)|$, in Fig. 7e shows that a higher value of m will result in a larger difference, meaning a more severe deformation in $R_{J_2 m}(\theta)$.



(a) Image J_1 (b) Image J_2

(c) $R_{J_1m}(\theta)$ (d) $R_{J_2m}(\theta)$ (e) $|R_{J_2m}(\theta) - R_{J_1m}(\theta)|$

Figure 7: Illustration of the sensitivity of the generic R-transform to sampling and quantization using an analytical triangle J_1 and its sampled and quantized version J_2 having a size of 100×100 pixels. A higher value of m will result in a larger difference between $R_{J_1m}(\theta)$ and $R_{J_2m}(\theta)$, meaning a more severe deformation in $R_{J_2m}(\theta)$.

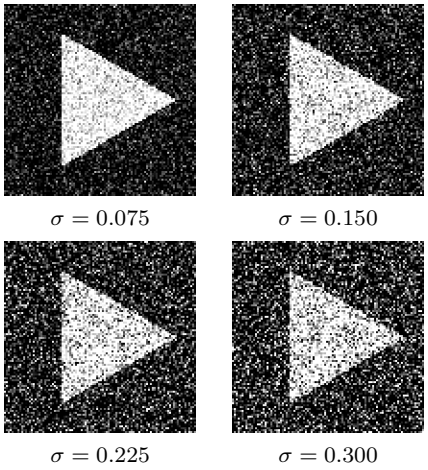
As the noise resulting from sampling and quantization is relatively small that may not demonstrate well the sensitivity of the generic R-transform to additive noise in general. A study has been carried out using noisy images generated from the image J_2 in Fig. 7b by adding white noise of different variances σ^2 to it. At each value of σ, 100 noisy images are generated for the computation of the average difference between the generic R-transform of J_2 and that of these noisy images:

$$\varepsilon_{R_{J_2}}(m,\sigma) = \frac{1}{|N_\sigma|} \sum_{J_{2i} \in N_\sigma} \int |R_{J_2m}(\theta) - R_{J_{2i}m}(\theta)| \, \mathrm{d}\theta,$$

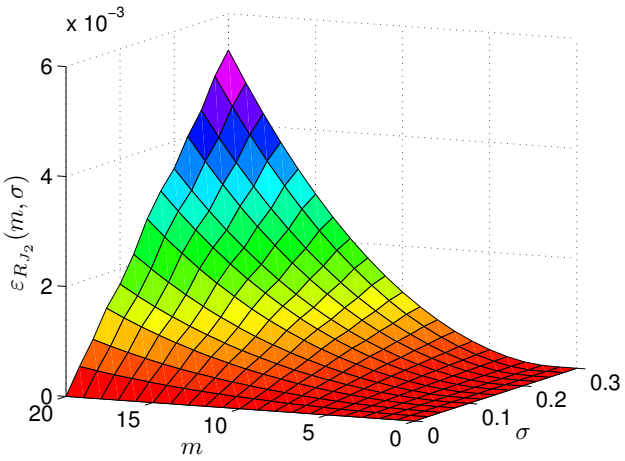
where N_σ is the subset of noisy images generated from J_2 having variance σ^2. Samples of noisy images generated from J_2 are given in Fig. 8a and the values of $\varepsilon_{R_{J_2}}(m,\sigma)$ are plotted in Fig. 8b. It is observed that $\varepsilon_{R_{J_2}}(m,\sigma)$ increases with both σ and m, meaning that a higher value of σ and/or m will result in a more severe deformation in $\varepsilon_{R_{J_2}}(m,\sigma)$. However, the increasing trend of $\varepsilon_{R_{J_2}}(m,\sigma)$ due to σ is different from that due to m: $\varepsilon_{R_{J_2}}(m,\sigma)$ tends to increase linearly with σ but exponentially with m. It is thus anticipated that the degradation in performance of the generic R-signature in invariant pattern recognition problems due to additive noise is linear with σ and exponential with m.

5 Robustness to noise

The Radon transform has been proven to be robust to additive noise due to the use of an integral function along straight lines in the spatial domain of pattern images [32]. For a grayscale image $f(x,y)$ of size $m \times n$ pixels having mean μ_s and variance σ_s^2 contaminated by additive white noise $\eta(x,y)$ of mean $\mu_n = 0$ and variance σ_n^2, the signal-to-noise ratio (SNR) of the noisy image $\hat{f}(x,y) = f(x,y) + \eta(x,y)$



(a) Samples of noisy images



(b) The average difference in the generic R-transform

Figure 8: The dependence on noise level σ and exponent m of the average difference between the generic R-transform of a noise-free image and that of its noisy versions. (a) Samples of noisy images generated from the image J_2 in Fig. 7b. (b) The average difference in the generic R-transform $\varepsilon_{R_{J_2}}(m,\sigma)$.

and its projection $\mathcal{R}_{\hat{f}}(\theta,\cdot)$ have been proven to have the relation:

$$\mathrm{SNR}_{\mathrm{proj}(\theta)} = \mathrm{SNR}_{\mathrm{image}} + \left(\frac{A(\theta)}{mn} - 1\right)\frac{\mu_s^2}{\sigma_n^2}, \tag{10}$$

where $A(\theta)$ is a constant depending both on θ, m, n. Similarly, when a binary image $f(x,y)$ having D percentage of pixels occupied by the shape region is contaminated by "salt & pepper" noise of flipping percentage d, the relation becomes

$$\frac{\mathrm{SNR}_{\mathrm{proj}(\theta)}}{\mathrm{SNR}_{\mathrm{image}}} = \frac{1 + D\left(\frac{A(\theta)}{mn} - 1\right)}{1 + d(1-2D)^2\left(\frac{A(\theta)}{mn} - 1\right)}. \tag{11}$$

As $\frac{A(\theta)}{mn} \simeq \sqrt{mn}$ then Eqs. (10) and (11) indicate a high increase in the value of SNR after projection, meaning that the Radon transform is very robust to additive noise.

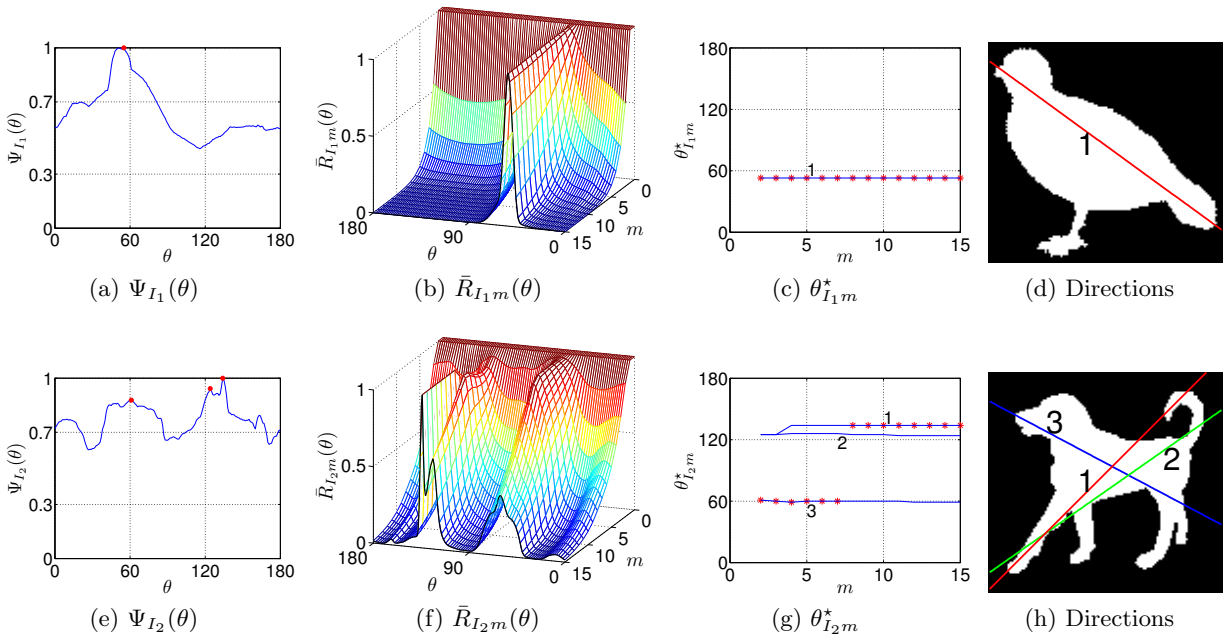
For the case of the generic R-transform, the use of exponentiation in its definition in Eq. (7) destroys the linearity and the uncorrelation between signal and noise. This in turn excludes the possibility of analyzing signal and noise separately and hinders the formulation of $\mathrm{SNR_R}$, SNR of the generic R-transform. Nevertheless, the following intuition suggests the sensibility of $\mathrm{SNR_R}$ on the exponent m. Consider a scalar signal f contaminated by noise η and assume that $m > 1$, the SNR before and after exponentiation are

$$\mathrm{SNR}_b = \frac{f^2}{\eta^2} \quad \text{and} \quad \mathrm{SNR}_a = \frac{f^2}{((f+\eta)^m - f^m)^2}$$

respectively. It is not difficult to see that

$$\frac{\mathrm{SNR}_a}{\mathrm{SNR}_b} = \frac{\eta^2}{((f+\eta)^m - f^m)^2}$$

decreases exponentially with the increase in m or, in other words, the decrease in SNR due to exponentiation depends exponentially on the exponent m. In the case of the generic R-transform, the integral after exponentiation has a smoothing property which alleviates the problem, especially for



(a) $\Psi_{I_1}(\theta)$ (b) $\bar{R}_{I_1 m}(\theta)$ (c) $\theta^\star_{I_1 m}$ (d) Directions

(e) $\Psi_{I_2}(\theta)$ (f) $\bar{R}_{I_2 m}(\theta)$ (g) $\theta^\star_{I_2 m}$ (h) Directions

Figure 9: Illustration of the ability of the generic R-transform and generic R-signature to encode pattern's dominant directions using two pattern images I_1, I_2 in Figs. 2a, 2b respectively. $\Psi_{I_i}(\theta)$ (first column) represents the highest accumulation of I_i along all $L(\theta, \cdot)$ and $\theta^\star_{I_i m}$ (third column) represents the ridges in the surface of $\bar{R}_{I_i m}(\theta)$ (second column). The fourth column shows the one (three) dominant directions of I_1 (I_2) corresponding to the one (three) ridges in $\bar{R}_{I_1 m}(\theta)$ ($\bar{R}_{I_2 m}(\theta)$).

big-sized images, similar to the smoothing property of the Radon transform's projection discussed above. However, the compensation is relatively small that the decrease in SNR still exists when m is reasonably large. This leads to a conclusion that a higher value of m will result in a lower value in SNR_R. An experimental support for this conclusion can be observed from Fig. 8b where the average difference $\varepsilon_{R_{J_2}}(m, \sigma)$ increases exponentially with m.

The robustness to noise of the generic R-signature has its roots not only from the noise-suppressing property of the Radon transform but also from the ability of the generic R-transform to encode dominant directions of patterns. Due to the exponentiation insides the integral in Eq. (7), the contribution of $\Psi_f(\theta) = \mathcal{R}_f(\theta, \rho^\star)$ with $\rho^\star = \text{argmax}_\rho \mathcal{R}_f(\theta, \rho)$ to $R_{fm}(\theta)$ increases as m increases, and furthermore:

$$\lim_{m \to +\infty} \frac{\Psi_f^m(\theta)}{R_{fm}(\theta)} = 1. \tag{12}$$

This means that, at a reasonable high value of m, $R_{fm}(\theta)$ represents the highest accumulation of $f(x, y)$ along all the lines $L(\theta, \cdot)$, which is similar to the "longest line" feature proposed in [14]. The highest accumulation profiles $\Psi_{I_1}(\theta)$ ($\Psi_{I_2}(\theta)$) of the pattern images I_1 (I_2) (normalized by $\max_\theta \Psi_{I_1}(\theta)$ ($\max_\theta \Psi_{I_2}(\theta)$) for better viewing) in Figs. 2a (2b) are plotted in Figs. 9a (9e) respectively.

Similarly, by denoting $\bar{R}_{fm}(\theta) = \frac{R_{fm}(\theta)}{R_{fm}(\theta^\star_{fm})}$ where $\theta^\star_{fm} = \text{argmax}_\theta R_{fm}(\theta)$ as a normalization of $R_{fm}(\theta)$, it is evident that $\bar{R}_{fm}(\theta)$ with $\theta \neq \theta^\star_{fm}$ decreases exponentially with the increase in m, and

$$\lim_{m \to +\infty} \bar{R}_{fm}(\theta) = \lim_{m \to +\infty} \frac{R_{fm}(\theta)}{R_{fm}(\theta^\star_{fm})} = \delta_{\theta\theta^\star_{fm}}, \tag{13}$$

where $\delta_{\theta\theta^\star_{fm}} = [\theta = \theta^\star_{fm}]$ is the Kronecker delta function. Thus, when m is reasonably high, the function $\bar{R}_{fm}(\theta)$ encodes only the direction $\theta^\star_{fm}$. Combining Eqs. (12) and (13) leads to a conclusion that

$\bar{R}_{fm}(\theta^{\star}_{fm})$ corresponds to the highest accumulation of $f(x,y)$ along all the lines $L(\cdot,\cdot)$ at a reasonable high value of m.

In real applications, Eqs. (12) and (13) do not hold as m does not have high enough value due to the sensitivity of the generic R-transform to quantization/sampling and additive noise as discussed in Subsection 4.2. However, there is an evolution in the profile of $\bar{R}_{fm}(\theta)$ as m increases, transforming a constant function ($m = 0, 1$) into the Kronecker delta function ($m = \infty$). During this process, the information encoded by $\bar{R}_{fm}(\theta)$ also changes, roughly from all directions to a single direction $\theta^{\star}_{fm}$, called the principal direction. The interpretation here is that the dominant directions of $f(x,y)$ are encoded at different levels, depending on m. Illustration of this evolution is depicted in Figs. 9b (9f) containing the plots of $\bar{R}_{I_1m}(\theta)$ ($\bar{R}_{I_2m}(\theta)$). The trace of the ridges in $\bar{R}_{I_1m}(\theta)$ ($\bar{R}_{I_2m}(\theta)$) is plotted in blue lines and the values of $\theta^{\star}_{I_1m}$ ($\theta^{\star}_{I_2m}$) are denoted by red asterisks in Figs. 9c (9g). It is observed that the one (three) ridges in the surface of $\bar{R}_{I_1m}(\theta)$ ($\bar{R}_{I_2m}(\theta)$) correspond to the one (three) local maxima of $\Psi_{I_1}(\theta)$ ($\Psi_{I_2}(\theta)$) in Figs. 9a (9e), which in turn represent the one (three) dominant directions of I_1 (I_2), as shown in Figs. 9d (9h). For the pattern image I_1, as there exists only one ridge, the principal direction always coincides with that ridge. In the case of the pattern image I_2, as m increases, the role of the three maxima in $\bar{R}_{I_2m}(\theta)$ interchanges along with a change in the principal direction from ridge 3 to ridge 1 at $m = 7$. However, dominant directions are still reflected in the profile of $\bar{R}_{I_2m}(\theta)$ as the local maxima of its three ridges.

When the pattern image $f(x,y)$ is contaminated by additive noise to be $\hat{f}(x,y)$, due to the noise-suppressing property of the Radon transform and the sensitivity of the generic R-transform to m, the difference in the dominant directions of $f(x,y)$ and $\hat{f}(x,y)$ is negligible when m is not too large. For this reason, $\bar{R}_{fm}(\theta)$ and the generic R-signature can be used to estimate the orientation of pattern images and to recognize noisy pattern images respectively when m is not too large. Combining with the dependance of the discriminatory power on m as discussed in Subsection 3.2, it can be concluded that the selected value of m is a compromise between two contradicting desires: a higher value is preferred for high discriminatory power whereas a lower one is for noise robustness. These theoretical analysis and observations will be supported by experimental evidences in the next section.

6 Experimental results

The effectiveness of the generic R-transform and generic R-signature has been demonstrated through two types of experiments in the following two subsections: one for the estimation of the principal direction of patterns and the other for invariant pattern recognition.

6.1 Fundamental direction estimation

The stability of the estimated principal direction of pattern images by the generic R-transform has been evaluated on the two datasets *OriA*, *OriB* of noisy images generated from the two pattern images I_1, I_2 in Figs. 2a, 2b respectively by adding white noise to them. These two images have been chosen as they are representatives of two classes of pattern images: I_1 belongs to the "easy" class while I_2 belongs to the "difficult" class due to the existence of the three ridges in the plot of $\bar{R}_{I_2m}(\theta)$ in Fig. 9f. Let SNR be the signal-to-noise ratio defined as

$$\mathrm{SNR} = \frac{\sum_{x,y} f^2(x,y)}{\sum_{x,y} \eta^2(x,y)},$$

where $f(x,y)$ is the noise-free image and $\eta(x,y)$ is the added white noise. Each dataset contains 600 noisy images of six possible values of SNR $= \{0.5, 1, 2, 4, 8, 16\}$, meaning 100 images for each SNR.

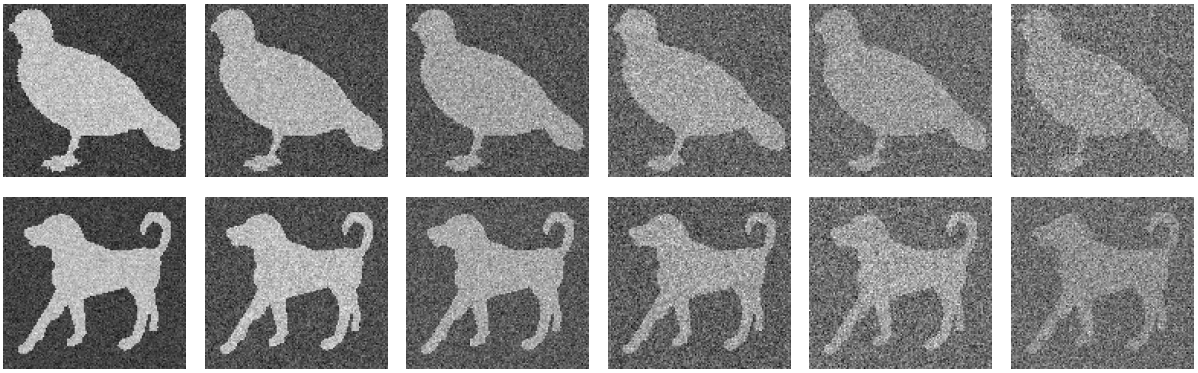


Figure 10: Example images of different SNR from the datasets *OriA* (top row) and *OriB* (bottom row) for principal direction estimation using the generic R-transform. These images are generated by adding white noise to the two pattern images I_1, I_2 in Figs. 2a, 2b respectively.

Example images of different SNR from these two datasets are given in Fig. 10: top row for *OriA* and bottom row for *OriB*. The generic R-transform of the noise-free images I_1, I_2 and all the noisy images in *OriA* and *OriB* has been computed along with their principal direction vector $\theta^\star_{fm}$ for evaluation.

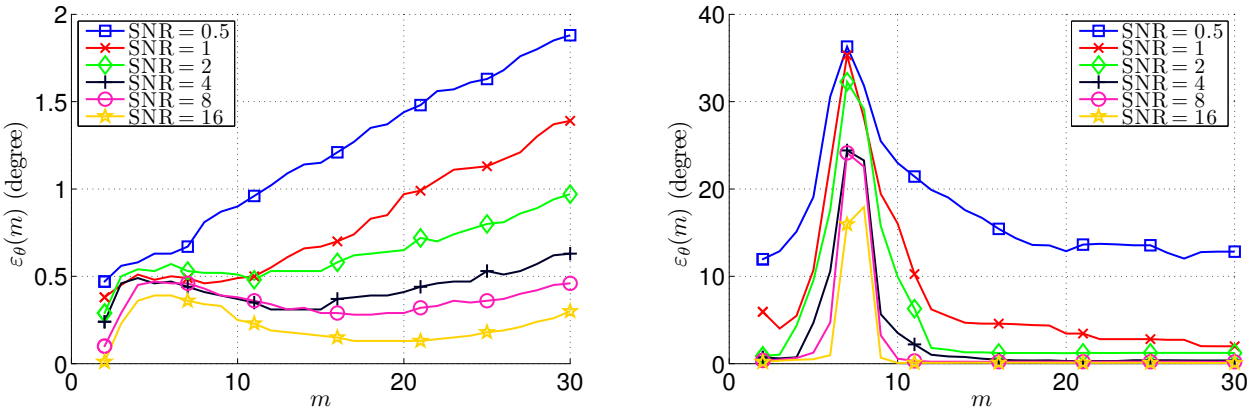
The adopted evaluation criteria is the average difference between the estimated principal direction of all noisy images of the same SNR in one dataset and their corresponding noise-free image as:

$$\varepsilon_\theta(m) = \frac{1}{|N_k|} \sum_{f_i \in N_k} |\theta^\star_{fm} - \theta^\star_{f_i m}| \quad \text{with } k \in \{0.5, 1, 2, 4, 8, 16\},$$

where f is a noise-free image and N_k is the subset of noisy images generated from f having SNR $= k$. This criteria measures statistically the effect of additive white noise on the accuracy of the estimated principal direction at different noise levels SNR and at different exponents m. Shown in Fig. 11 are the plots of the computed $\varepsilon_\theta(m)$ at different SNR for the two datasets *OriA* and *OriB*, it is observed that:

- *OriA*: The values of $\varepsilon_\theta(m)$ shown in Fig. 11a are generally small ($\varepsilon_\theta(m) < 2^\circ$ for $m < 30$), demonstrating the stability of the estimated principal direction $\theta^\star_{I_1 m}$. Additionally, the values of $\varepsilon_\theta(m)$ increases with the increase in m; high accuracy ($\varepsilon_\theta(m) < 0.5^\circ$) is obtained when $m \simeq 2$ for all SNR. Explanation for this, besides the noise-suppressing property of the Radon transform, comes from the sensitivity of the generic R-transform to m as discussed in Section 5. When m is small, the smoothing property of integral in Eq. (7) allows $\mathcal{R}_f(\theta, \cdot)$ to participate in $\bar{R}_{fm}(\theta)$ and thus further reduces the effect of additive white noise. When m increases, this smoothing property gradually disappears as the role of $\mathcal{R}_f(\theta, \rho^\star)$ in $\bar{R}_{fm}(\theta)$ gradually dominates.

- *OriB*: The values of $\varepsilon_\theta(m)$ shown in Fig. 11b have a different trend from those in Fig. 11a; $\varepsilon_\theta(m)$ has its peak value at $m = 7$ for almost all SNR. This is due to the existence of the three ridges in the plot of $\bar{R}_{I_2 m}(\theta)$ as given in Fig. 9f: the role of encoding the principal direction of the pattern image I_2 changes from ridge 3 to ridge 1 at $m = 7$ (shown in Fig. 9g) while, in the presence of additive noise, the changing point is not always at $m = 7$. Additionally, as seen in Fig. 9f, the ridge encoding the principal direction at each m does not have a decisive role, the two remaining ridges always have inference, making the estimated principal direction vulnerable to additive noise and resulting in high value of $\varepsilon_\theta(m) \simeq 12^\circ$ when SNR $= 0.5$. However, when the noise is weak (SNR ≥ 2), high accuracy is still obtained (for example, $\varepsilon_\theta(m) < 2^\circ$ at $m = 30$).



(a) *OriA* (b) *OriB*

Figure 11: The average difference in the estimated principal direction $\varepsilon_\theta(m)$ between the noise-free image and the noisy images of different SNR of the two datasets *OriA* (a) and *OriB* (b). A small value of $\varepsilon_\theta(m)$ means that the principal direction estimated by the generic R-transform is very stable under noise perturbation.

These observations lead to a conclusion that the estimated principal direction of pattern images by the generic R-transform is very stable under noise perturbation when m is not too large, and by a simple extension, similar conclusion for patterns' dominant directions could also be reached. They provides experimental evidences for theoretical arguments in Section 5.

6.2 Invariant pattern recognition

Two experiments on grayscale and binary image datasets have been carried out to demonstrate the robustness of the proposed generic R-signature in invariant pattern recognition problems. The robustness of the proposed descriptor to additive white noise is first demonstrated by using two sets of datasets, generated from images of 26 Latin characters and from the COIL-20 dataset [33] by adding white noise of different levels to them. Secondly, the proposed descriptor is computed on a set of datasets generated from the UMD Logo dataset [34] by adding "salt & pepper" noise of different levels to its images. The aim of these experiments is to demonstrate the robustness to "salt & pepper" noise of the proposed descriptor. Thus, the first experiment deals with grayscale patterns and the second one with binary patterns.

The proposed generic R-signature is compared with angular radial transform (ART) [35], generic Fourier descriptor (GFD) [6], Zernike moments [36], Radon 1D Fourier–Mellin transforms (RFM) [13], and Radon 2D Fourier–Mellin transforms (R2DFM) [12]. Except for the RFM, all other comparison descriptors need normalizations in order to be invariant to RST transformations. Moreover, the RFM and R2DFM descriptors are also defined on the Radon transform. These descriptors are selected because they are commonly used and have good reported performance.

Similarity measure

For any two pattern images I and J represented invariantly by $FR_{Im}(\xi)$ and $FR_{Jm}(\xi)$ respectively, their measure of similarity is defined as the ℓ_2-norm distance between their descriptor as

$$\mathrm{dist}(I, J) = \|FR_{Im}(\xi) - FR_{Jm}(\xi)\|_2. \tag{14}$$

The computation of $\mathrm{dist}(I, J)$ is simple and fast, permitting the generic R-signature to be used in pattern matching problems with large-size datasets. More sophisticated distances like the weighted Euclidean distance [36] could be used to reduce the dominance of some of the coefficients in the generic R-signature. However, as small-valued coefficients usually correspond to high frequency components, meaning that they are more sensitive to additive noise and sampling/quantization effect, balancing the coefficient contributions thus reduces the performance of the descriptor in noisy environment. The performance degradation resulting from coefficient weighting has been observed from some preliminary experiments. Moreover, due to the orthogonality in the basis of the discrete Fourier transform, there is no correlation among coefficients of the generic R-signature and thus the Mahalanobis distance [37], if employed, reduces to the weighted Euclidean distance.

Evaluation criterion

The criterion used for comparison among descriptors is the precision–recall curve defined in information retrieval context [38]. Denoting:

- *retrieved images*: the list of images produced by a matching process for a query image.
- *relevant images*: the list of all images in the dataset being in the same category with the query image.

Then, precision is defined as the fraction of retrieved images that are relevant to the search:

$$\text{Precision} = \frac{|\{\text{relevant images}\} \cap \{\text{retrieved images}\}|}{|\{\text{retrieved images}\}|},$$

and recall is defined as the percent of all relevant images that is returned by the search:

$$\text{Recall} = \frac{|\{\text{relevant images}\} \cap \{\text{retrieved images}\}|}{|\{\text{relevant images}\}|}.$$

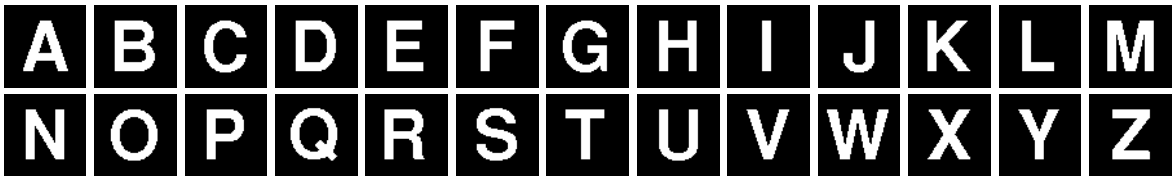
In computing the precision–recall curve for each dataset, in the experiment, each of the images in the dataset is used as a model to which all the images in the dataset are compared/matched with. The matching is realized using the similarity measure defined in Eq. (14). The obtained matching results are then sorted, or ranked, for the determination of the nth nearest matches for each model.

6.2.1 Grayscale pattern images

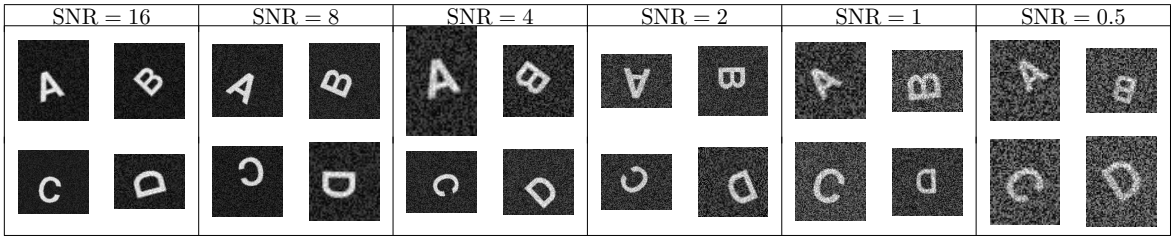
The performance of the proposed generic R-signature has been first tested on grayscale noisy images to demonstrate its robustness to additive white noise. Two experiments have been carried out on two different sets of datasets:

- *ExpA*: The first set of six alphabet datasets has been generated from images of 26 Latin characters as shown in Fig. 12a. Each of these six datasets has 260 images of 26 categories, each category contains 10 images.
- *ExpB*: The second set of six object datasets has been generated from 20 object images from the COIL-20 dataset [33] as shown in Fig. 13a. Each of these six datasets has 220 images of 20 categories, each category contains 11 images.

The main characteristic that differentiates *ExpA* and *ExpB*, besides the semantic content of their images, is the number of intensity levels in the original images: character images have only two levels of



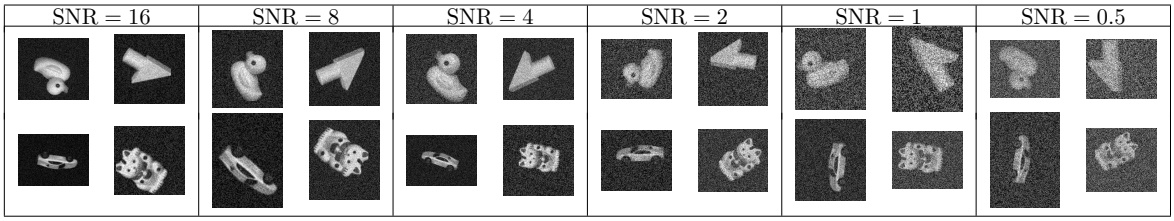
(a) Noise-free images



(b) Examples of noisy images

Figure 12: (a) Images of 26 Latin characters of size 64×64 pixels in Arial bold font used to generate the six alphabet datasets. (b) Sample images from the six alphabet datasets generated from the first four character images with six possible values of SNR $= \{16, 8, 4, 2, 1, 0.5\}$, corresponding to the six datasets.

(a) Noise-free images



(b) Examples of noisy images

Figure 13: (a) Twenty object images from the COIL-20 dataset used to generate the six object datasets. (b) Sample images from the six object datasets generated from the four object images with six possible values of SNR $= \{16, 8, 4, 2, 1, 0.5\}$, corresponding to the six datasets.

intensity whereas object images have multi-level intensity. Noisy grayscale images are generated from the corresponding noise-free images by randomly scaling, rotating, translating and then adding white noise to them. The value of SNR for each dataset is kept constant and, for each experiment, SNR has six possible values $\{0.5, 1, 2, 4, 8, 16\}$, corresponding to the six datasets. Some example images from the six datasets in the two experiments are given in Figs. 12b–13b.

Figs. 14 (15) provide the precision–recall curves of the generic R-signature computed on the six character (object) datasets of *ExpA* (*ExpB*). In these figures, at a specific value of m in the horizontal axis, there is a precision–recall curve with recall and precision rates illustrated as the ordinate and the color of the grid points having abscissa m. It is observed that the performance of the generic R-signature varies according to m. As m increases from 0.2 to 10 and except for the singularity $m = 1$,

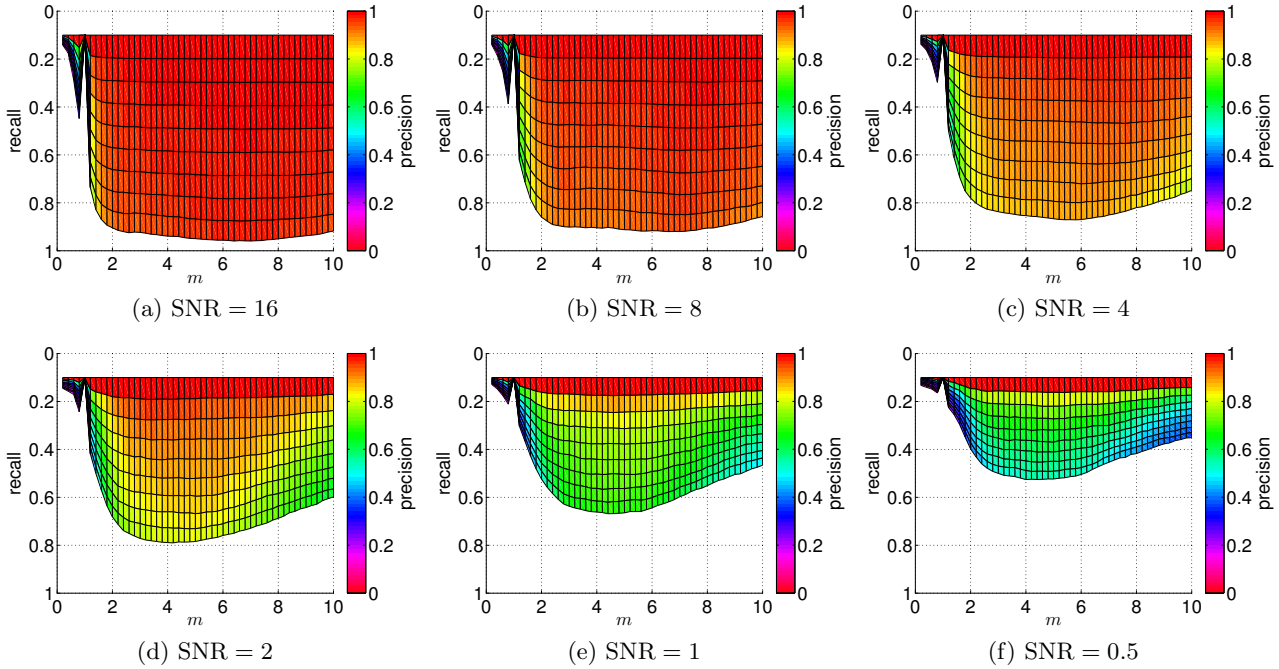
the precision–recall curve, when plotted in the traditional 2D Cartesian coordinate system with recall and precision as the abscissa and ordinate respectively, goes upwards till a certain value of m and then downwards, meaning an increase and then a decrease in performance of the generic R-signature. In general, the peak in performance is obtained at $m \simeq 5$ ($\simeq 3.2$) for *ExpA* (*ExpB*), leading to a conclusion that the selected value of m to have best performance does not depend on the level of noise. The increase in performance at low value of m agrees with the increase in discriminatory power, which results from exploiting variation in the intersection of the pattern image with parallel lines, as discussed in Subsection 3.2. Additionally, the decrease in performance at high value of m agrees with the discussion on the sensitivity of the generic R-signature to additive noise in Section 5. Furthermore, as SNR increases, the performance of the generic R-signature generally deteriorates at each value of m, which agrees with the dependance of the generic R-signature on noise level presented in Subsection 4.2. However, the deterioration speed is slower at $m \simeq 5$ ($\simeq 3.2$) for *ExpA* (*ExpB*), meaning a robustness of the generic R-signature to additive white noise.

Due to the existence of a class of signatures, their combination has also been investigated to see if it leads to possible gains in performance. For simplicity and for the reasons that will be clear later, two R-signatures of exponents m_1 and m_2 has been combined, $FR_{Im_1m_2}(\xi) = \left[FR_{Im_1}(\xi), FR_{Im_2}(\xi)\right]$, to be used as the invariant descriptor for the pattern image I. Figs. 16 (17) provide the accuracy of the combined R-signature $FR_{Im_1m_2}(\xi)$ on the six character (object) datasets of *ExpA* (*ExpB*). In these figures, at a specific value of m_1 and m_2 in the horizontal and vertical axes, the accuracy is illustrated as the color of the grid point (m_1, m_2). It can be seen that the color pattern of these figures is symmetric with respect to the minor diagonal and a change in (m_1, m_2) generally leads to a change in the color, meaning that the performance of $FR_{Im_1m_2}(\xi)$ varies according to (m_1, m_2). As m_1 and m_2 are interchangeable and should be different to avoid duplicate, assuming that $m_1 < m_2$. The peak in performance is then obtained at $m_1 = 2.6, m_2 = 5.2$ ($m_1 = 2.4, m_2 = 3.8$) for *ExpA* (*ExpB*). Note from these values of m_1 and m_2 that one is smaller and the other is larger than m. These relations among the selected values of exponents have the following possible explanations:

- m_1 and m_2 should be separated enough to make use of the difference in the discriminatory information contained in $FR_{Im_1}(\xi)$ and $FR_{Im_2}(\xi)$.
- m_1 and m_2 should be close to m so that $FR_{Im_1}(\xi)$ and $FR_{Im_2}(\xi)$ individually has high discriminatory power, similar to that of $FR_{Im}(\xi)$.

Comparison of the proposed generic R-signature and RFM descriptor with ART, GFD, Zernike, and R2DFM descriptors on these noisy datasets have been performed and the obtained results are given in Figs. 18 and 19 respectively. In this comparison, besides the conventional value of 2, the value of the exponent m has been selected reflecting the relatively best performance of the generic R-signature on *ExpA* ($m = 5$) and *ExpB* ($m = 3.2$). In addition, the combined R-signature has also been used in the comparison with $m_1 = 2.6, m_2 = 5.2$ ($m_1 = 2.4, m_2 = 3.8$) for *ExpA* (*ExpB*). It is observed from these sets of figures that:

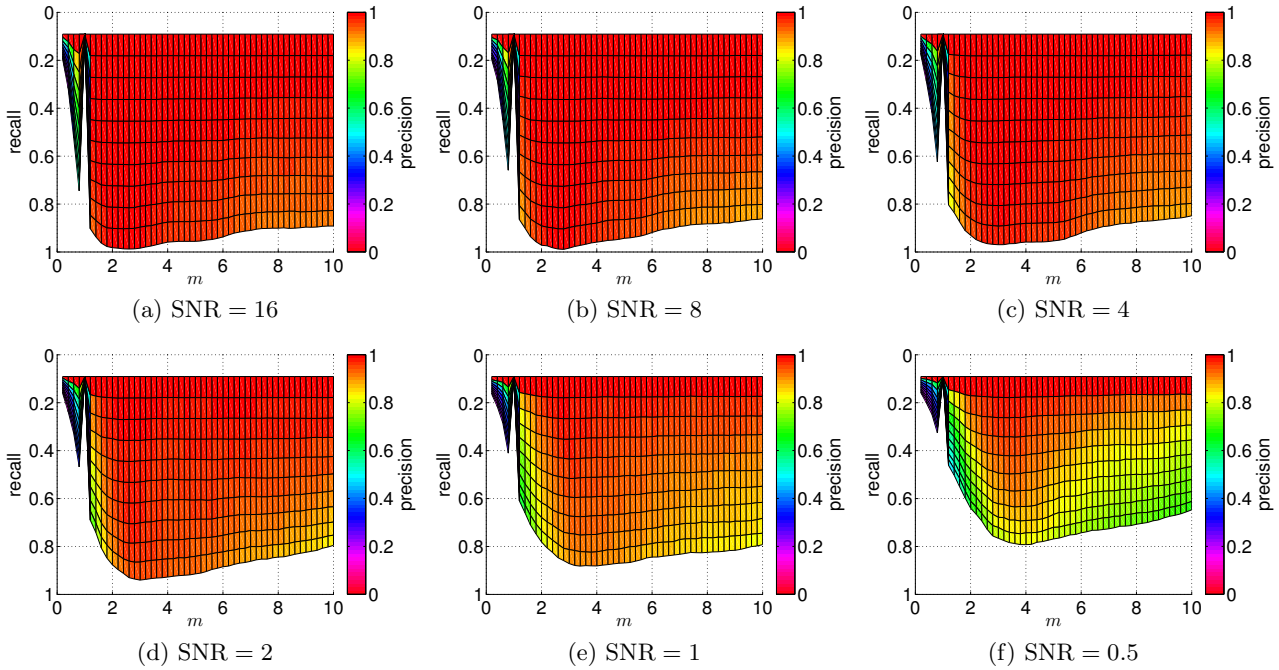
- ART, GFD, Zernike, and R2DFM descriptors are not robust to additive white noise at all, their performance is similarly poor for different levels of noise.
- There is a substantial gain in performance of the generic R-signature from the conventional R-signature ($m = 2$) when an appropriate value of the exponent m is used.
- As SNR decreases (the images get noisier), the precision–recall curve of the generic R-signature and RFM descriptor generally moves downwards, having comparable and good performance.



(a) SNR = 16 (b) SNR = 8 (c) SNR = 4

(d) SNR = 2 (e) SNR = 1 (f) SNR = 0.5

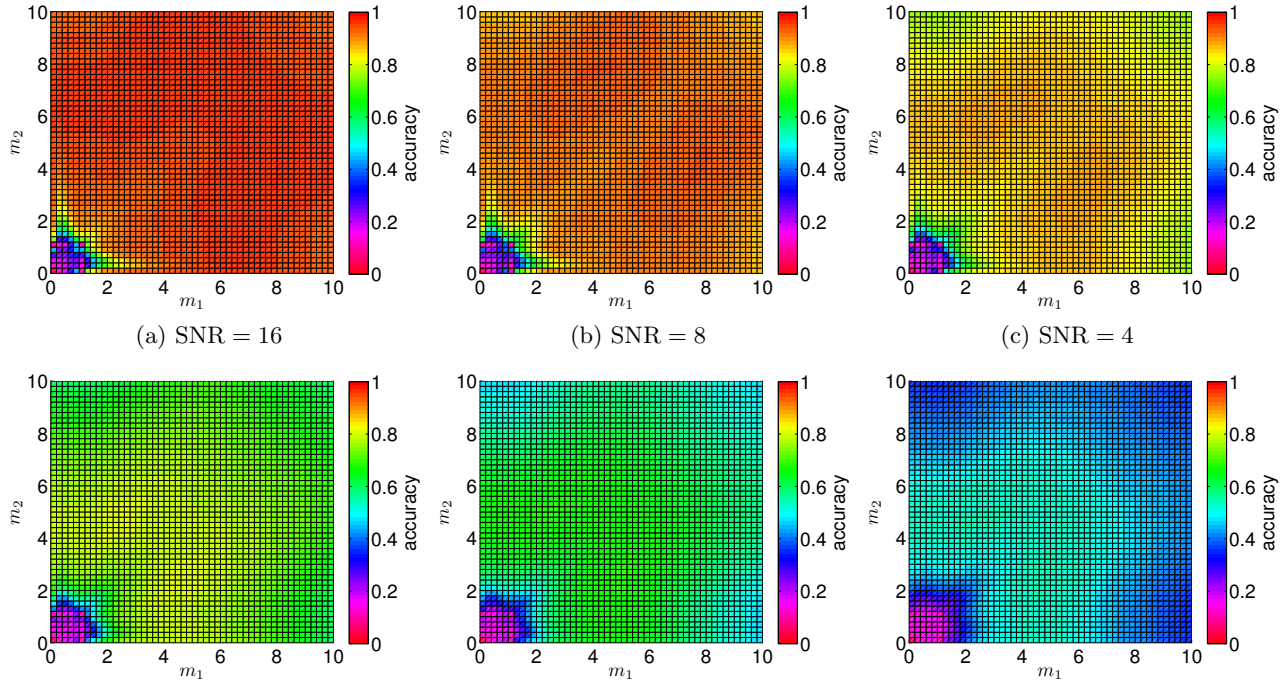
Figure 14: Precision–recall curves of the generic R-signature on the six alphabet datasets at different values of m. In each of these figures, at a specific value of m in the horizontal axis, there is a precision–recall curve with recall and precision rates illustrated as the ordinate and the color of the grid points having abscissa m.



(a) SNR = 16 (b) SNR = 8 (c) SNR = 4

(d) SNR = 2 (e) SNR = 1 (f) SNR = 0.5

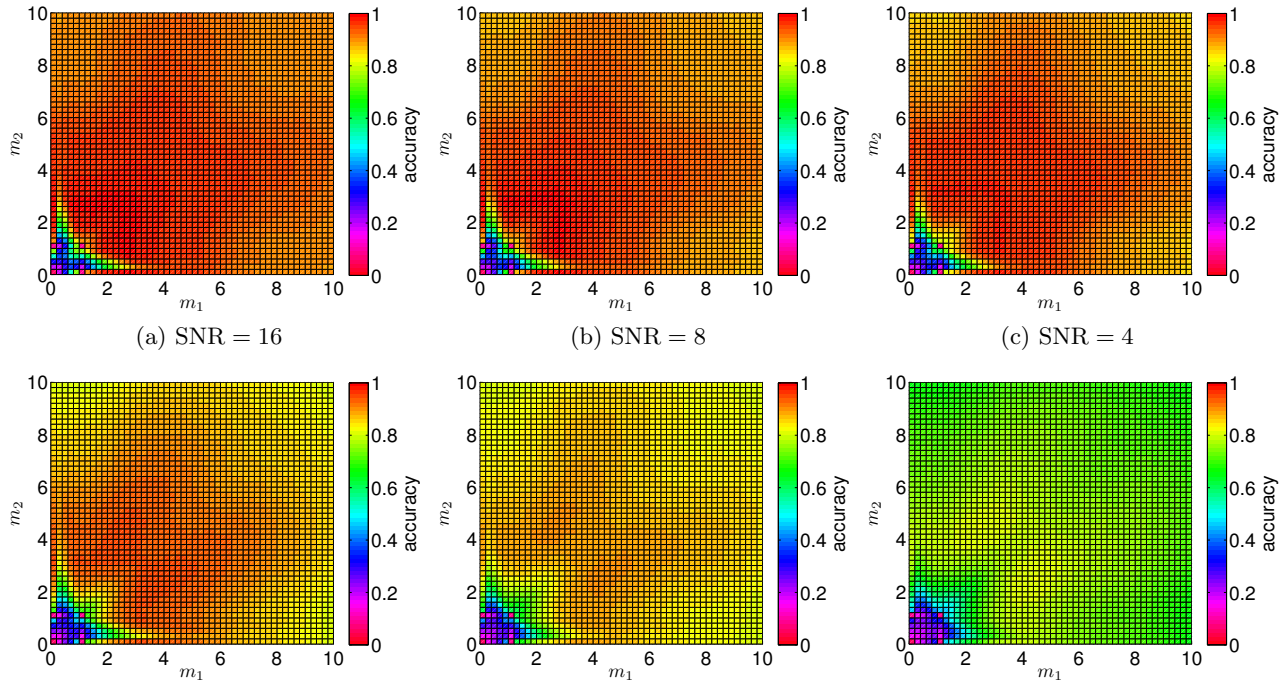
Figure 15: Precision–recall curves of the generic R-signature on the six object datasets at different values of m. In each of these figures, at a specific value of m in the horizontal axis, there is a precision–recall curve with recall and precision rates illustrated as the ordinate and the color of the grid points having abscissa m.



(a) SNR = 16 (b) SNR = 8 (c) SNR = 4

(d) SNR = 2 (e) SNR = 1 (f) SNR = 0.5

Figure 16: The accuracy of the generic R-signature on the six alphabet datasets at different values of (m_1, m_2). In each of these figures, at a specific value of (m_1, m_2), the accuracy is denoted as the color of the grid point having abscissa m_1 and ordinate m_2.



(a) SNR = 16 (b) SNR = 8 (c) SNR = 4

(d) SNR = 2 (e) SNR = 1 (f) SNR = 0.5

Figure 17: The accuracy of the generic R-signature on the six object datasets at different values of (m_1, m_2). In each of these figures, at a specific value of (m_1, m_2), the accuracy is denoted as the color of the grid point having abscissa m_1 and ordinate m_2.

Their performance is nearly perfect when the noise is weak (SNR $= 16, 8, 4$), demonstrating their robustness to additive noise.

- The combined R-signature does perform better than the single one. However, the gain in performance is very small and negligible.

It is thus can be concluded that the proposed generic R-signature is more robust to additive white noise than the comparison ART, GFD, Zernike, and R2DFM descriptors and has comparable performance to the RFM descriptor on grayscale noisy datasets. This provides empirical evidences for the analytical results developed in Section 5. Disappointing gain resulting from combination of generic R-signatures can be explained by their "similar nature" in describing pattern images. The only difference among them is the difference in the value of the exponent m, meaning a difference in the exploitation of the variation in $\mathcal{R}_f(\theta, \cdot)$ (Subsection 3.2). This conclusion can also be generalized that even the combined R-signature is composed of more generic R-signatures, the performance gain is still not noteworthy. Moreover, poor performance of ART, GFD, Zernike, and R2DFM descriptors has its root in the required normalizations in their computation and can be explained as:

- To have invariance to translation, the origin of the polar coordinate system needs to be located at the centroid of the pattern. In the presence of noise, position of the centroid is shifted arbitrarily according to the actual noise.

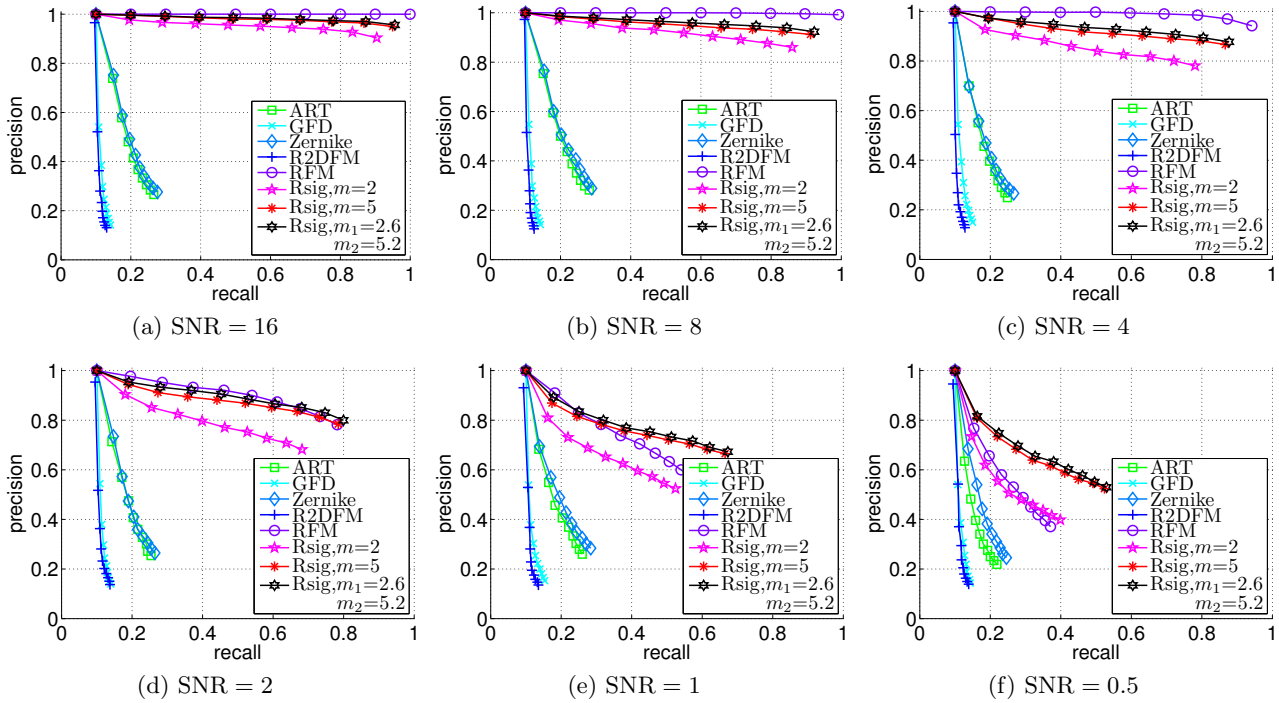
- To have invariance to scaling, the radial axis is normalized by the distance from the origin of the polar coordinate system to the farthest pattern's point. In the presence of noise, this farthest point might not belong to the actual pattern but the noise.

Furthermore, it is also evident from the above two sets of experiments that the performance of the generic R-signature and RFM descriptor is better on *ExpB* than on *ExpA* at each value of SNR, leading to a conclusion that the proposed descriptor performs better on multi-level than on two-level grayscale images. Possible explanation for this comes from the Radon transform values: multi-level images tend to have more variation in their Radon transform than that of two-level ones. In addition, recall from Subsection 3.2 that the role of m is to exploit the variation in constant-θ slices of the Radon image with more variation usually leads to higher discriminatory power. Thus at the same value of m, the generic R-signature of multi-level images contains more discriminatory power than that of two-level ones.

6.2.2 Binary pattern images

The robustness of the proposed generic R-signature to additive "salt & pepper" noise is demonstrated using a set of six logo datasets generated from the first 25 logo images of the UMD Logo dataset [34] as shown in Fig. 20b. Each of these six logo datasets has 275 images of 25 categories, each category contains 11 images generated by randomly scaling, rotating, translating the original corresponding logo image and then adding "salt & pepper" noise to it. Let d be the percentage of pixels flipped by the noise, the value of d for each generated dataset is kept constant and d has six possible values, ranging from 0 to 0.1 with increment of 0.02 corresponding to the six datasets. The first dataset with $d = 0$ is actually a noiseless dataset; its use is intended for checking the invariant properties of the proposed and comparison descriptors. The values of d of the other five noisy datasets make up an arithmetic progression with a common difference of 0.02. These five datasets are, therefore, used to evaluate the robustness of the proposed and comparison descriptors at incrementing levels of additive "salt & pepper" noise. Some sample images from these six datasets are given in Fig. 20b.

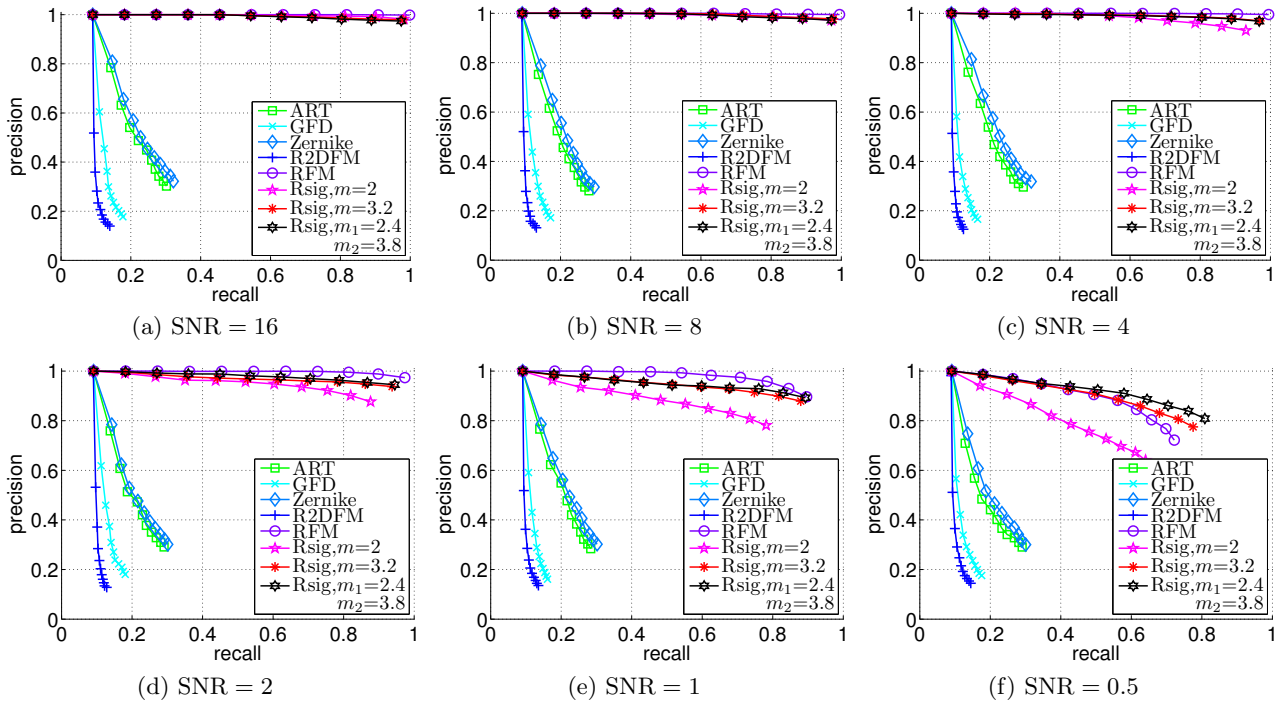
Fig. 21 provides the precision–recall curves of the generic R-signature on the six logo datasets. The evolution of these curves according to m has similar trend with that on the six alphabet and object



(a) SNR = 16 (b) SNR = 8 (c) SNR = 4

(d) SNR = 2 (e) SNR = 1 (f) SNR = 0.5

Figure 18: Precision–recall curves of comparison descriptors on the six alphabet datasets. ART, GFD, Zernike, and R2DFM descriptors are not robust to noise, their curves are similarly poor at different values of SNR while generic R-signature and RFM descriptor are. As SNR decreases, the curves of generic R-signature and RFM descriptor generally move downwards.

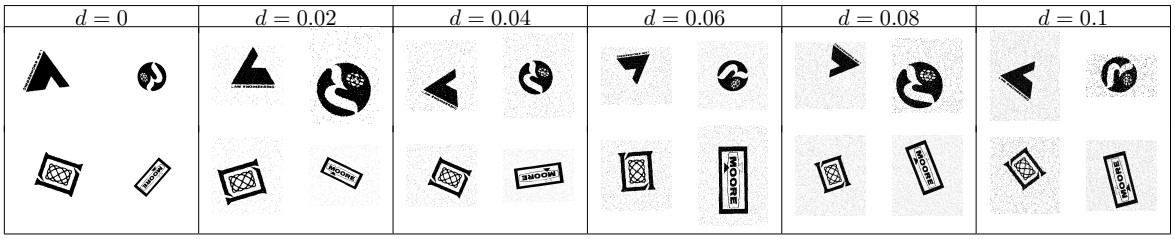


(a) SNR = 16 (b) SNR = 8 (c) SNR = 4

(d) SNR = 2 (e) SNR = 1 (f) SNR = 0.5

Figure 19: Precision–recall curves of comparison descriptors on the six object datasets. ART, GFD, Zernike, and R2DFM descriptors are not robust to noise, their curves are similarly poor at different values of SNR while generic R-signature and RFM descriptor are. As SNR decreases, the curves of generic R-signature and RFM descriptor generally move downwards.

(a) Noise-free images



(b) Examples of noisy images

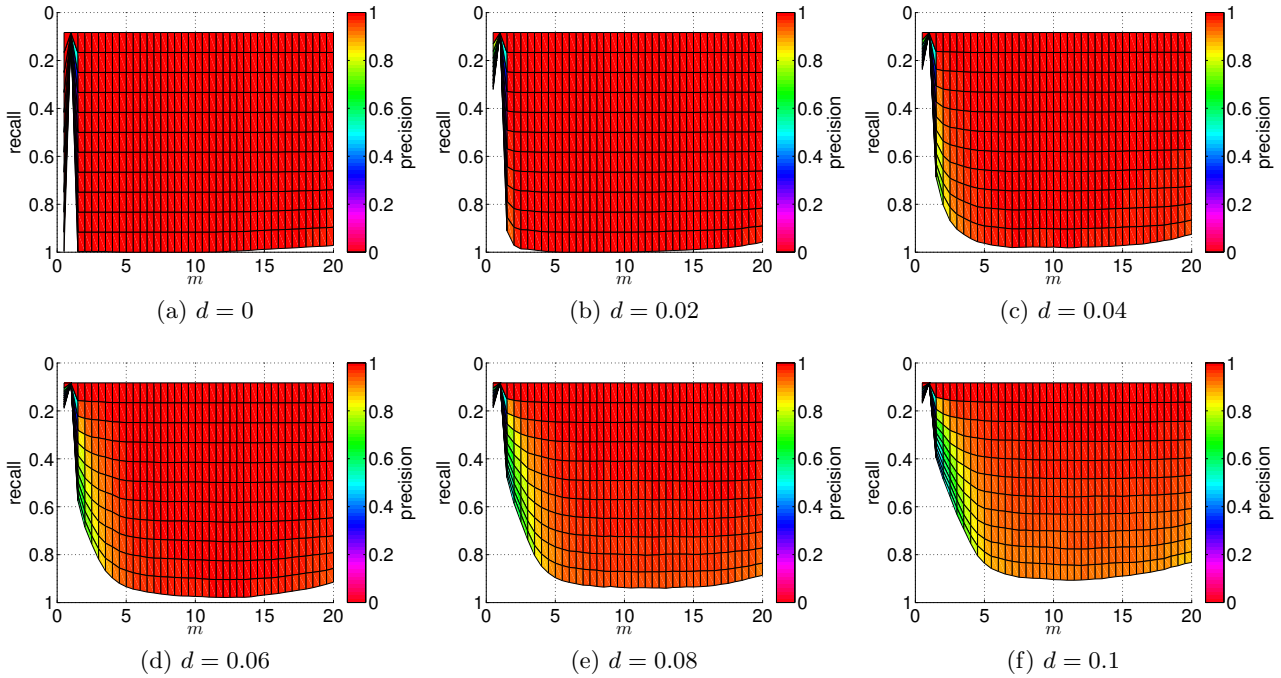
Figure 20: (a) Twenty-five logo images from the UMD Logo dataset used to generate the six logo datasets. (b) Sample images from the six logo datasets generated from the first four logo images with six possible values of $d = \{0, 0.02, 0.04, 0.06, 0.08, 0.1\}$, corresponding to the six datasets.

datasets in Figs. 14 and 15 respectively. That is, except for the singularity at $m = 1$, an increase then a decrease in performance are observed, agreeing with the discussions in Subsection 3.2 and Section 5 respectively. The peak in performance is obtained at $m \in [9, 13]$ and, as d increases, the performance of the generic R-signature generally deteriorates at each value of m, which agrees with the dependance of the generic R-signature on noise level presented in Subsection 4.2. However, the deterioration speed is slow at $m \in [9, 13]$, meaning a robustness of the generic R-signature to additive "salt & pepper" noise.

The accuracy of the combined R-signature $FR_{Im_1m_2}(\xi)$ on the six logo datasets is given in Fig. 22. Similar to the accuracy plots in Figs. 16 (17), it can be seen that the color pattern of these figures is symmetric with respect to the minor diagonal and a change in (m_1, m_2) generally leads to a change in the color, meaning that the performance of $FR_{Im_1m_2}(\xi)$ varies according to (m_1, m_2). The peak in performance is obtained at $m_1 = 5.5, m_2 = 11.5$. It should be noted again from these values of m_1 and m_2 that one is smaller and the other is larger than m.

The proposed generic R-signature is again compared with ART, GFD, Zernike, RFM, and R2DFM descriptors using these six datasets and the computed precision–recall curves of these descriptors are depicted in Fig. 23. In this comparison, the value of m is fixed at 11 for the single signature and $m_1 = 5.5, m_2 = 11.5$ for the combined one. For the noiseless dataset with $d = 0$ (Fig. 23a), all shape descriptors have ideal performance, demonstrating that the proposed descriptor is totally invariant to both translation, rotation, and scaling. When $d \neq 0$ (Fig. 23b–23f), deterioration appears in the performance of all descriptors and their precision–recall curve moves downwards. However, the impact of d on those curves differs from one descriptor to another. Among the other descriptors, the proposed descriptor has the best performance for all the five noisy datasets while ART and Zernike descriptors have similarly worse performance. It is also observed that:

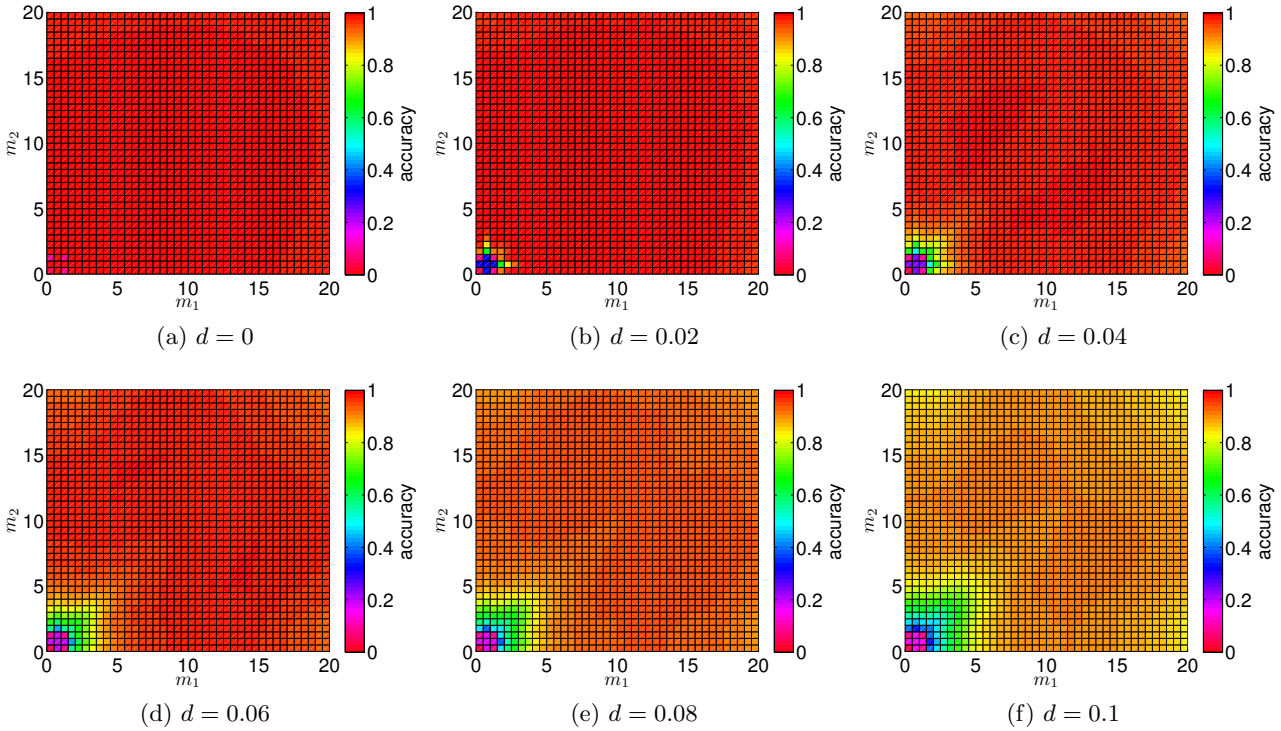
- As d increases, the curve of all the descriptors generally moves downwards.
- ART and Zernike descriptors are not robust to noise at all, their performance is similarly poor for different levels of noise.
- GFD have more resistance to noise than ART and Zernike because its curves are pushed away from the ideal curve (when $d = 0$) with distance which increases along with the increase in d.



(a) $d = 0$ (b) $d = 0.02$ (c) $d = 0.04$

(d) $d = 0.06$ (e) $d = 0.08$ (f) $d = 0.1$

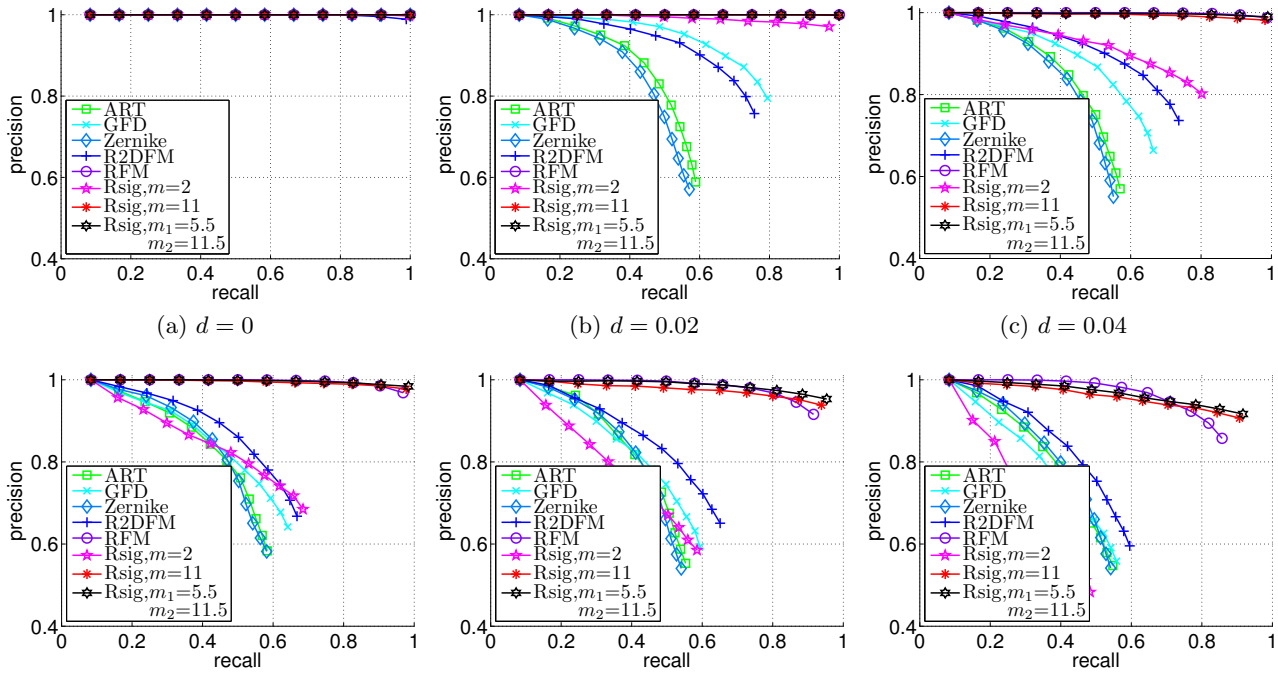
Figure 21: Precision–recall curves of the generic R-signature on the six logo datasets at different values of m. In each of these figures, at a specific value of m in the horizontal axis, there is a precision–recall curve with recall and precision rates illustrated as the ordinate and the color of the grid points having abscissa m.



(a) $d = 0$ (b) $d = 0.02$ (c) $d = 0.04$

(d) $d = 0.06$ (e) $d = 0.08$ (f) $d = 0.1$

Figure 22: The accuracy of the generic R-signature on the six logo datasets at different values of (m_1, m_2). In each of these figures, at a specific value of (m_1, m_2), the accuracy is denoted as the color of the grid point having abscissa m_1 and ordinate m_2.



(a) $d = 0$ (b) $d = 0.02$ (c) $d = 0.04$

(d) $d = 0.06$ (e) $d = 0.08$ (f) $d = 0.1$

Figure 23: Precision–recall curves of comparison descriptors on the six logo datasets. For the noiseless dataset $d = 0$ (a), all pattern descriptors have ideal performance. When $d \neq 0$ (b)–(f), deterioration appears in the performance of all descriptors and their curve moves downwards. However, the impact of d on those curves differs from one descriptor to another.

However, the resistance of GFD is weaker than that of descriptors defined on the Radon transform.

- Among the three Radon transform-based descriptors, the shift in the curve of the generic R-signature and RFM descriptor is similar and more regular than that of R2DFM.
- There is a substantial gain in performance of the generic R-signature from the conventional R-signature ($m = 2$) when an appropriate value of the exponent m is used.
- The gain in performance obtained by combining the generic R-signatures is small and negligible.

The above observations lead to a conclusion that the proposed generic R-signature is more robust to additive "salt & pepper" noise than the comparison ART, GFD, Zernike, and R2DFM descriptors and has comparable performance to the RFM descriptor on binary noisy datasets. This provides empirical evidences for the analytical results developed in Section 5. Explanations for small gain due to R-signature combination and poor performance of ART, GFD, Zernike, and R2DFM descriptors on binary noisy datasets are similar to those given in the previous subsection for grayscale noisy datasets.

Performance of the proposed generic R-signature on deformation datasets such as MPEG-7 has also been evaluated. The obtained results are not promising and cannot compete with methods designed specifically to tolerate shape deformation such as the one in [39] and references therein. Poor performance of the proposed method is due to its nature: the use of transforms in its definition in order to have invariance to RST transformations reduces drastically its capability to tolerate shape deformation. However, it should be noted that:

- The proposed method is not intended nor designed to work solely with binary patterns. It is

designed, instead, to work with grayscale patterns under RST transformations allowing a certain level of additive noise.

- Methods such as in [39] cannot work with noisy patterns and need "clean" images for feature extraction, which the proposed method does not.
- Due to the use of the Euclidean distance, matching two patterns by means of their generic R-signature is simple and fast, allowing it to be used in large-scale dataset.

7 Conclusions

This paper has provided a generic view on popular Radon transform-based transform and descriptor, the R-transform and R-signature. The generic R-signature brings in a class of descriptors having the beneficial properties of the conventional R-signature while spatially describing patterns at all directions and at different levels. This generalization gives more flexibility in their definition and, more importantly, the generic R-signature has been proven to be robust to additive noise and demonstrated its superiority over existing invariant pattern descriptors on grayscale and binary noisy datasets. The proper value of the exponent m, the only parameter of this generalization which does not depend on the level of noise, is constrained by the two contradicting desires (a higher value is preferred for high discriminatory power whereas a lower one is for noise robustness) and depends on the content of images. As the discriminatory power results from exploiting variation in the intersection of the pattern image with parallel lines, it is anticipated that a pattern image that has less variation in its spatial domain will require a higher value of m for best performance. Evidences are $m = 3.2, 5, 11$ for the three types of datasets used in experiments: object, alphabet, and logo. Moreover, due to the blunt maxima in the accuracy curves of the generic R-signature, a small deviation of the selected value of m from the best choice has almost no effect on the performance. Additionally, the generic R-transform has the ability of encoding patterns' dominant directions, opening up some potential applications like patterns' orientation estimation or document image skew correction. Future works will explore further applications of the generic R-transform in pattern recognition.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable and insightful comments and suggestions that help to improve the quality of the paper.

A Proof of the properties of the generic R-transform

- *Periodicity*: Using property $P3$ of the Radon transform:

$$\begin{aligned} R_{fm}(\theta) &= \int_{-\infty}^{\infty} \mathcal{R}_f^m(\theta, \rho)\, d\rho = \int_{-\infty}^{\infty} \mathcal{R}_f^m(\theta \pm \pi, -\rho)\, d\rho = -\int_{\infty}^{-\infty} \mathcal{R}_f^m(\theta \pm \pi, \upsilon)\, d\upsilon \\ &= \int_{-\infty}^{\infty} \mathcal{R}_f^m(\theta \pm \pi, \upsilon)\, d\upsilon = R_{fm}(\theta \pm \pi), \end{aligned}$$

  using $\upsilon = -\rho$.

- *Translation*: Assuming $f(x, y)$ is translated by a vector $\vec{u} = (x_0, y_0)$, using property $P4$ of the Radon transform:

$$\begin{aligned} R_{f'm}(\theta) &= \int_{-\infty}^{\infty} \mathcal{R}^m_{f'}(\theta, \rho)\, d\rho = \int_{-\infty}^{\infty} \mathcal{R}^m_f(\theta, \rho - x_0 \cos\theta - y_0 \sin\theta)\, d\rho \\ &= \int_{-\infty}^{\infty} \mathcal{R}^m_f(\theta, \upsilon)\, d\upsilon = R_{fm}(\theta), \end{aligned}$$

  using $\upsilon = \rho - x_0 \cos\theta - y_0 \sin\theta$.

- *Rotation*: Assuming $f(x, y)$ is rotated by an angle θ_0, using property $P5$ of the Radon transform:

$$R_{f'm}(\theta) = \int_{-\infty}^{\infty} \mathcal{R}^m_{f'}(\theta, \rho)\, d\rho = \int_{-\infty}^{\infty} \mathcal{R}^m_f(\theta + \theta_0, \rho)\, d\rho = R_{fm}(\theta + \theta_0).$$

- *Scaling*: Assuming $f(x, y)$ is scaled by a factor α, using property $P6$ of the Radon transform:

$$\begin{aligned} R_{f'm}(\theta) &= \int_{-\infty}^{\infty} \mathcal{R}^m_{f'}(\theta, \rho)\, d\rho = \int_{-\infty}^{\infty} \frac{1}{\alpha^m} \mathcal{R}^m_f(\theta, \alpha\rho)\, d\rho \\ &= \frac{1}{\alpha^{m+1}} \int_{-\infty}^{\infty} \mathcal{R}^m_f(\theta, \upsilon)\, d\upsilon = \frac{1}{\alpha^{m+1}} R_{fm}(\theta). \end{aligned}$$

References

[1] J. Wood, "Invariant pattern recognition: a review," *Pattern Recognition*, vol. 29, no. 1, pp. 1–17, 1996. 1

[2] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern Recognition*, vol. 37, no. 1, pp. 1–19, 2004. 1

[3] K. B. Howell, "Fourier transforms," in *Transforms and Applications Handbook*, 3rd ed., A. D. Poularikas, Ed. CRC Press, 2010, ch. 2. 1

[4] J. Bertrand, P. Bertrand, and J. P. Ovarlez, "Mellin transform," in *Transforms and Applications Handbook*, 3rd ed., A. D. Poularikas, Ed. CRC Press, 2010, ch. 12. 1

[5] Y.-N. Hsu, H. H. Arsenault, and G. April, "Rotation-invariant digital pattern recognition using circular harmonic expansion," *Applied Optics*, vol. 21, no. 22, pp. 4012–4015, 1982. 1

[6] D. Zhang and G. Lu, "Shape-based image retrieval using generic Fourier descriptor," *Signal Processing: Image Communication*, vol. 17, no. 10, pp. 825–848, 2002. 1, 6.2

[7] M. R. Teague, "Image analysis via the general theory of moments," *Journal of the Optical Society of America*, vol. 70, no. 8, pp. 920–930, 1980. 1

[8] A. Averbuch, R. Coifman, D. Donoho, M. Israeli, and J. Waldén, "Fast Slant Stack: A Notion of Radon Transform for Data on A Cartesian Grid Which Is Rapidly Computable, Algebraically Exact, Geometrically Faithful, and Invertible," Stanford University, Tech. Rep., 2001. 1

[9] H. Hjouj and D. W. Kammler, "Identification of reflected, scaled, translated, and rotated objects from their Radon projections," *IEEE Transactions on Image Processing*, vol. 17, no. 3, pp. 301–310, 2008. 1, 2.1

[10] S. Tabbone, L. Wendling, and J.-P. Salmon, "A new shape descriptor defined on the Radon transform," *Computer Vision and Image Understanding*, vol. 102, no. 1, pp. 42–51, 2006. 1, 2.2, 2.2, 3, 3.4

[11] N. Nacereddine, S. Tabbone, D. Ziou, and L. Hamami, "Shape-based image retrieval using a new descriptor based on the Radon and wavelet transforms," in *Proceedings of the 20th International Conference on Pattern Recognition*, 2010, pp. 1997–2000. 1

[12] X. Wang, B. Xiao, J.-F. Ma, and X.-L. Bi, "Scaling and rotation invariant analysis approach to object recognition based on Radon and Fourier–Mellin transforms," *Pattern Recognition*, vol. 40, no. 12, pp. 3503–3508, 2007. 1, 3.4, 6.2

[13] T. V. Hoang and S. Tabbone, "A geometric invariant shape descriptor based on the Radon, Fourier, and Mellin transforms," in *Proceedings of the 20th International Conference on Pattern Recognition*, 2010, pp. 2085–2088. 1, 3.1, 3.4, 6.2

[14] Y. W. Chen and Y. Q. Chen, "Invariant description and retrieval of planar shapes using Radon composite features," *IEEE Transactions on Signal Processing*, vol. 56, no. 10-1, pp. 4762–4771, 2008. 1, 5

[15] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972. 1

[16] V. F. Leavers and J. F. Boyce, "The Radon transform and its application to shape parametrization in machine vision," *Image and Vision Computing*, vol. 5, no. 2, pp. 161–166, 1987. 1

[17] V. F. Leavers, "Use of the Radon transform as a method of extracting information about shape in two dimensions," *Image and Vision Computing*, vol. 10, no. 2, pp. 99–107, 1992. 1

[18] ——, "Use of the two-dimensional Radon transform to generate a taxonomy of shape for the characterization of abrasive powder particles," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1411–1423, 2000. 1

[19] A. Kadyrov and M. Petrou, "The trace transform and its applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 8, pp. 811–828, 2001. 1

[20] J. Li, S. K. Zhou, and R. Chellappa, "Appearance modeling using a geometric transform," *IEEE Transactions on Image Processing*, vol. 18, no. 4, pp. 889–902, 2009. 1

[21] O. Ramos-Terrades, E. Valveny, and S. Tabbone, "Optimal classifiers fusion in a non-Bayesian probabilistic framework," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 9, pp. 1630–1644, 2009. 1

[22] R. Souvenir and K. Parrigan, "Viewpoint manifolds for action recognition," *EURASIP Journal on Image and Video Processing*, vol. 2009, pp. 1–13, 2009. 1

[23] Y. Wang, K. Huang, and T. Tan, "Human activity recognition based on R transform," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2007. 1

[24] T. V. Hoang and S. Tabbone, "Text extraction from graphical document images using sparse representation," in *Proceedings of the 9th International Workshop on Document Analysis Systems*, 2010, pp. 143–150. 1

[25] ——, "Generic r-transform for invariant pattern representation," in *Proceedings of the 9th International Workshop on Content-Based Multimedia Indexing*, 2011, pp. 157–162. 1

[26] S. R. Deans, *The Radon Transform and Some of Its Applications*. Krieger Publishing Company, 1993. 2.1

[27] T. P. Wallace and P. A. Wintz, "An efficient three-dimensional aircraft recognition algorithm using normalized Fourier descriptors," *Computer Graphics and Image Processing*, vol. 13, no. 2, pp. 99–126, 1980. 2.2

[28] G. Borgefors, "Distance transformations in digital images," *Computer Vision, Graphics, and Image Processing*, vol. 34, no. 3, pp. 344–371, 1986. 3

[29] W. A. Götz and H. J. Druckmüller, "A fast digital Radon transform – an efficient means for evaluating the Hough transform," *Pattern Recognition*, vol. 29, no. 4, pp. 711–718, 1996. 3.4

[30] M. L. Brady, "A fast discrete approximation algorithm for the Radon transform," *SIAM Journal on Computing*, vol. 27, no. 1, pp. 107–119, 1998. 3.4

[31] K. Jafari-Khouzani and H. Soltanian-Zadeh, "Rotation-invariant multiresolution texture analysis using Radon and wavelet transforms," *IEEE Transactions on Image Processing*, vol. 14, no. 6, pp. 783–795, 2005. 4.2

[32] T. V. Hoang and S. Tabbone, "Invariant pattern recognition using the rfm descriptor," *Pattern Recognition*, vol. 45, no. 1, pp. 271–284, 2012. 5

[33] S. A. Nene, S. K. Nayar, and H. Murase, "Columbia Object Image Library (COIL-20)," Department of Computer Science, Columbia University, Tech. Rep. CUCS-005-96, 1996. 6.2, 6.2.1

[34] D. S. Doermann, E. Rivlin, and I. Weiss, "Applying algebraic and differential invariants for logo recognition," *Machine Vision and Applications*, vol. 9, no. 2, pp. 73–86, 1996. 6.2, 6.2.2

[35] M. Bober, F. Preteux, and W.-Y. Y. Kim, "Shape descriptors," in *Introduction to MPEG 7: Multimedia Content Description Language*, B. S. Manjunat, P. Salembier, and T. Sikora, Eds. John Wiley & Sons, 2002, ch. 15, pp. 231–260. 6.2

[36] A. Khotanzad and Y. H. Hong, "Invariant image recognition by Zernike moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 5, pp. 489–497, 1990. 6.2, 6.2

[37] P. C. Mahalanobis, "On the generalised distance in statistics," *Proceedings of the National Institute of Sciences of India*, vol. 2, no. 1, pp. 49–55, 1936. 6.2

[38] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*. The ACM Press, 1999. 6.2

[39] X. Bai, X. Yang, L. J. Latecki, W. Liu, and Z. Tu, "Learning context sensitive shape similarity by graph transduction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 861–874, 2010. 6.2.2